The Role of Molecular Oxygen in the Epigenetic Regulation of Cellular Differentiation

Burr, Simon Jonathan

Awarding institution:
King's College London

The Role of Molecular Oxygen in the Epigenetic Regulation of Cellular Differentiation

Simon Burr

February 2018

Faculty of Life Sciences and Medicine
Department of Cardiology
British Heart Foundation Centre of Research Excellence
King's College London

Submitted for the Degree of Doctor of Philosophy

Supervisors: Dr. Alison C. Brewer & Prof. Giovanni E. Mann

“I began to realize how important it was to be an enthusiast in life. He taught me that if you are interested in something, no matter what it is, go at it at full speed ahead. Embrace it with both arms, hug it, love it, and above all become passionate about it. Lukewarm is no good. Hot is no good either. White hot and passionate is the only thing to be.”

Roald Dahl – My Uncle Oswald

Table of Contents

Table of Figures

Table of Tables

Acknowledgments

I would firstly like to thank my primary supervisor Dr. Alison Brewer for her consistent guidance and support throughout my PhD project. I am especially appreciative for the opportunity to explore and develop my own research ideas, enabling myself to develop as a research scientist. I particularly admired her drive and passion, which kept me motivated when faced with frustration and technical difficulty. I would have not achieved the success or had such a positive PhD experience without her mentorship.

Extended thanks go to Anna Caldwell for mass-spectrometry work, Mei Chong for her NMR expertise, as well as Sucharitha Balu and Matthew Arno for NGS sequencing help. Additional colleagues that have taught me all things science, provided life advice and I regard as good friends include Rajesh Mistry, Dan Martin, Matthew Hancock and Stephen Metcalf. Other lab heroes include Matteo Beretta, Thomas Murray, Brodie Quine, Stephen Grey, Valeria León Kropf, Temo Barwari, Izaj Rahman and Giulia Emanuelli. On a more personal level I want to like to give a special thank you to Natalie Savage, not only for being the perfect tea making buddy and accompanying me on chocolate seeking expeditions, but for her continued encouragement and support. I will always be grateful to her for putting up with me during my times of thesis writing stress and for the fond memories of our fun trips away. Furthermore, my time at King's College London wouldn't have been as enjoyable without my fellow 4-year British Heart Foundation MRes/PhD programme year group: Karen Frudd, Hannah Lewis and Phoebe Kitscha.

I would like to thank the support of my family, especially my parents and grandparents for believing in me, being there to listen to my troubles and their consistent guidance. Additional thanks go out to my friends for always being on hand should I need them, particularly Matthew Price, James Collins, Matthew Wade, Kathryn Davidson, Bethany Jackson, Pippa Clark, Kerri Little and Marie Collard.

Lastly, I would like to thank the British Heart Foundation for their sponsorship and making my research possible.

Declaration

I declare that I am the sole author of this thesis and that the work contained within it is my own.

Simon Burr

Abstract

The developing mammalian embryo is subjected to low O_2 tensions, which owing to the short diffusion distance of O_2 may form a gradient within the early embryo, and thus may function as a developmental morphogen. It was observed that culture of mouse embryonic stem cells (mESCs) in 1% O_2, compared to atmospheric [O_2], acted to skew differentiation. The cause of this effect was studied with respect to alterations in chromatin structure, *via* regulation of epigenetic modifications. The ten-eleven translocation (Tet) demethylase enzymes, Tet1/2/3, oxidise methylated DNA to form 5-hydroxymethylcytosine (5-hmC); a stable epigenetic mark that is implicated in developmental events. The mechanism(s) which underlie the regulation of their function(s) are unknown. However, intriguingly the catalytic activity of these enzymes is mediated *via* a conserved O_2-dependent hydroxylase domain. Therefore, it was hypothesised that the Tet enzymes may be differentially regulated by [O_2] to influence cellular specification and thus contribute to asymmetry within the developing embryo. Here it was identified that Tet1 is the most likely isoform to be inhibited by O_2 tensions deemed physiologically relevant during embryogenesis. Further, differentiating mESCs displayed a transient, Tet1-mediated, O_2-dependent burst of hydroxylation. This hydroxylation was predominantly targeted to a CG rich region of a Tet3 promoter, driving expression of a truncated protein isoform that lacks a CXXC DNA binding domain. This promoter region was shown to associate with Tet1, and also both the repressive (H3K27me3) and activating (H3K4me3) histone marks, characteristic of a bivalent promoter. Here it was also confirmed that 2 distinct Tet3 protein isoforms are expressed in differentiated mESCs, which were found to have differential transcriptional regulation and are thus likely to serve distinct cellular functions. It is suggested that Tet1 activity, in part determined by [O_2] within the early embryo, may regulate Tet3 expression spatially to control cell fate decisions.

Abbreviations

2-HG: 2-hydroxyglutarate
2-OG: 2-oxoglutarate
2-OGDO: 2-oxoglutarate dependent dioxygenases
3'UTR: 3’-untranslated regions
5-hmC: 5-hydroxymethylcytosine
5-mC: 5-methylcytosine
5'RACE: 5’Rapid Amplfication of cDNA ends
AA: Ascorbic acid
ADCRs: ATP-dependent remodelling complexes
AMD: Age-related macular degeneration
AML: Acute myeloid leukemia
ATP: Adenosine triphosphate
BER: Base-excision repair
bp: Base pairs
BSO: Buthionine sulfoxime
C: Cytosine
ChIP: Chromatin immunoprecipitation
CMML: Chronic myleomonocytic leukemia
CpG: CG dinucleotides
DFO: Desferrioxamine
Dgcr8: DiGeorge syndrome critical region 8
DHA: Dehydroascorbic acid
DMOG: Dimethyloxalyglycine
DMSO: Dimethyl sulfoxide
Dnmt: DNA methyltransferase
DSBH: Distorted double stranded β-helix
E: Embryonic day
EBs: Embryoid bodies
EMT: Epithelial to mesenchymal transition
Epi: Epiblast
EPO: Erythropoietin
ExE: Extraembryonic ectoderm
FBS: Fetal bovine serum
FGF: Fibroblast growth factor
FH: Fumarate hydratase
FIH: Factor inhibiting HIF
GlcNAc: β-N-acetylglucosamine
GSH: Glutathione
GSK: Glycogen synthase kinase-3
Gulo: L-gulonolactone-γ-oxidase
H3K27me2: Histone H3 lysine 27 dimethylation
H3K27me3: Histone H3 lysine 27 trimethylation

H3K4me3: Histone H3 lysine 4 trimethylation
HCF1: Host cell factor 1
HEK cells: Human embryonic kidney cells 293
hESCs: human embryonic stem cells
HIF: Hypoxic inducible factors
HMT: histone methyltransferases
hMedip-seq/PCR: 5-hmC-DNA immunoprecipitation-sequencing/PCR
HPLC: High pressure liquid chromatography
HRE: Hypoxic response elements
HSPC: Hematopoietic stem/progenitor cells
ICM: Inner cell mass
IDH: Isocitrate dehydrogenase
IP: Immunoprecipitation
iPSCs: induced pluripotent stem cells
JmjC: Jumonji C
K: Lysine
KD: Knockdown
KDM: Histone demethylases
KO: KnockOut™ (Serum replacement)
L-gulonolactone-γ-oxidase: Gulo
LB: Lysogeny broth
LIF: Leukemia inhibitory factor
lncRNA: long ncRNAs
LSD: Lysine specific demethylase
MAPK: Mitogen-activated protein kinase
MBD: Methyl-CpG-binding domain
MEFs: Mouse embryonic fibroblasts
mESCs: mouse embryonic stem cells
miRNAs: microRNAs
MPN: Myeloproliferative neoplasms
mRNA: messenger RNA
ncRNA: non-coding RNAs
NPCs: Neural progenitor cells
O-GlcNac: O-linked β-N-acetylglucosamine
OGT: O-GlcNac transferase
OxPhos: Oxidative phosphorylation
PBS/T: Phosphate buffered saline/ containing 0.1% tween-20
PcGs: Polycomb group of proteins
PCR: Polymerase chain reaction
PGCs: Primordial germ cells
PHD: Prolyl hydroxylase domain
piRNAs: piwi-interacting RNAs
Prc1: Polycomb repressor complex 1
Prc2: Polycomb repressor 2
PrE: Primitive endoderm

Precursor-miRNAs: pre-miRNAs
PS: Primitive streak
pVHL: von Hippel-Lindau
QPCR: Quantitative polymerase chain reaction
RA: Retinoic acid
RISC: RNA induced silencing complex
RNAse: ribonuclease III
RPE: Retinal pigment epithelial
SAH: S-adenosylhomocysteine
SAM: S-adensylmethionine
SDG: Succinate dehydrogenase
SEM: Standard error of the mean
shRNA: short hairpin RNA
siRNA: short interfering RNAs
T: Thymine
TBS/T: Tris-buffered salt solution/ containing 0.1% tween-20
TCA: Tricarboxylic acid cycle
TDG: Thymine DNA glycosylase
TE: Trophectoderm
Tet: Ten-eleven translocation enzymes
TFs: Transcription factors
TKO: Triple Tet silenced
TLC: Thin layer chromatography
U: Uracil
VE: Visceral endoderm
VEGF: Vascular endothelial growth factor

Chapter 1: General Introduction

1.1 *Regenerative Medicine: Stem cell therapies*

Advancements in medicine have enabled effective treatments against a plethora of pathological conditions enabling a better quality of life and prolonged survival for much of the first world population. However, many of these treatments either slow progression or relieve symptoms of a disease rather than acting directing to repair the damaged tissue or organ. Thus, in this new age of modern medicine there is a paradigm shift away from the development of molecular pharmaceutical agents towards methods of regeneration. Regenerative medicine encompasses any method to replace or regenerate cells, tissues or organs to restore normal biological function[1]. A major arm of regenerative medicine is cell-based therapies. Cell therapy involves the introduction of live cells either alone, or in combination with a scaffold (tissue engineering), directly to the patient, to repair the target tissue or organ[2]. The concept of cell therapy dates back to the first successful blood transfusions in the 19th century[3], and has now evolved to be the focus of many research laboratories and the pharmaceutical industry worldwide.

Cell therapy has the potential to be applied to a variety of medical conditions, including heart failure, neurodegenerative diseases and spinal cord injury. For example, heart failure, which can result from myocardial infarction, culminates in the loss of cardiomyocytes and hence heart hemodynamic function. This loss of contractile machinery results in an inability to pump blood sufficiently to sustain physiological function and can often prove fatal. This is in part due to the fact that the mammalian heart, compared to amphibians and zebrafish, has a limited regenerative capacity, and thus is unable to effectively facilitate *de novo* generation of a functional myocardium[4]. Therefore, the replacement of cells, with the aim of restoring biological function, provides an exciting avenue of scientific exploration. Significant advances in this field have led to some, albeit limited, success in clinical studies[5,6], however perhaps the most profound and convincing proof-of-concept studies for cellular therapy in cardiac regeneration was conducted by Chong, J *et al* (2014)[7]. Here it was shown human embryonic stem cell (hESC) derived cardiomyocytes could substantially engraft and remuscularise the non-human primate heart in a model of myocardial ischemia[7].

Mesenchymal stem cells, a type of adult stem cell, obtained initially from bone marrow, but also found in adipose tissue and umbilical cord blood[8], offer a cell source for therapeutic applications. Specifically, these multipotent cells can differentiate into osteoblasts and chondrocytes and have thus been assessed for their ability to treat conditions such as osteoarthritis[9]. However, ESCs have been at forefront of much research owing to their unique pluripotency properties, allowing them to differentiate into

cells of all 3 embryonic germ layers[10]. 3 methods are commonly used to promote ESC differentiation: the formation of 3-dimensional cell aggregates named embryoid bodies (EBs), the culture of cells as monolayers on extracellular matrix proteins and culture of cells directly on a stroma-cell supportive layer[11]. This differentiation can be directed to form for instance cardiomyocytes and neurons[12,13,14,15,16,17], with the aim to exogenously transplant such cells to the target site for therapeutic use (see Figure 1:1 for current cell therapies undergoing clinical testing).

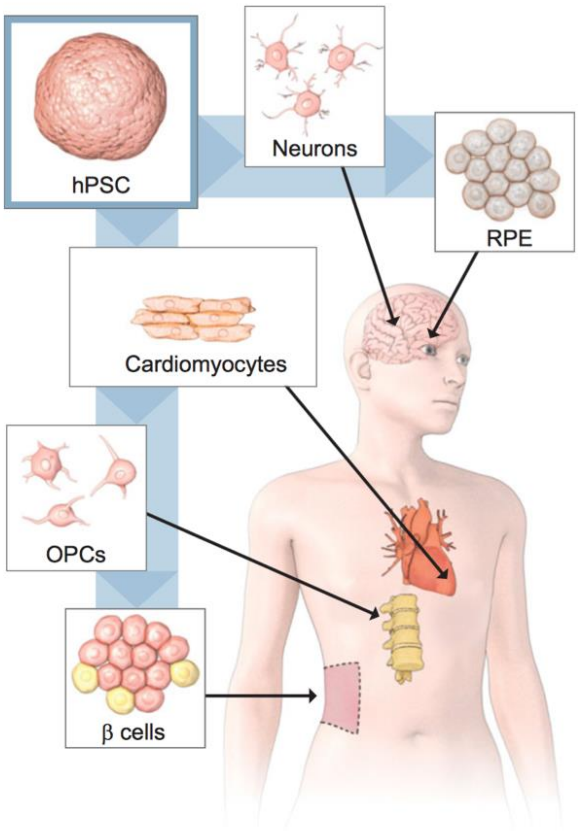


Figure 1:1 Clinical applications for human pluripotent stem cells in regenerative medicine. The use of human pluripotent stem cell derived therapeutics undergoing clinical testing include neurons (Parkinson's disease), retinal pigment epithelial cells (retinal degenerative diseases), cardiomyocytes (heart disease), oligodendrocytes progenitor cells (OPCs) (spinal cord injury) and β-islet cells (β cells) (type 1 diabetes). Image taken from Thies, R and Murry, C. Development. 2015; 142:3077-3084[6].

The use of ESCs are regarded as a powerful tool for regeneration strategies, however the development of cell therapies has been impeded by ethical constraints and immuno-incompatibility[18]. However, the development of reprogramming technology, which is the ability to restore pluripotency to somatic cells through ectopic expression of 4 key

transcription factors (TFs) (Oct4, Sox2, Klf4 and c-Myc), may have overcome these issues[19]. The generation of these patient-specific induced pluripotent stem cells (iPSCs) can be used in cell therapy and disease modelling. This type of cell therapy enables autologous generation of terminally differentiated cells for transplantation back into the patient, thereby reducing the requirement of immune suppression[20]. An example of how such a treatment has been used clinically is in the treatment of age-related macular degeneration. An iPSC-derived transplantation of retinal pigment epithelial cells revealed, albeit in a single patient, functional integration of these autologous cells, which halted visual deterioration[21]. However, in a contrasting report intravitreal injections of autologous adipose-tissue derived stem cells in 3 AMD patients resulted in loss of vision[22]. Thus collectively knowledge regarding how stem cells can be properly differentiated into desired cell types is paramount to ensure safe and efficacious therapies[23]. In addition to cell based therapies, such cells can be used for disease modelling, enabling the recapitulation of a cellular phenotype that occurs in a specific patient, which for example may have a genetic defect. These cells can subsequently be used for drug screening and toxicology studies with the aim of treating similar patients and thus paving the way for personalised medicine. Currently, various tissue-specific iPSC derived models have been generated, including neurological[24], and cardiovascular[25], which can provide a useful tool to investigate pathophysiology[26].

Despite the huge applicability of cell-based regenerative therapies there are concerns that are hindering their therapeutic use. Aside from the requirement to generate adequate numbers of cells for the clinic, one of the most fundamental residing challenges is the ability to generate homogenous fully differentiated cell populations, which eliminates the risk of teratoma formation, a characteristic pluripotent trait that occurs in both ESCs and iPSCs upon transplantation[27]. Teratomas are characterised by a disorganised arrangement of differentiated tissues originating from a pluripotent cell, and are histologically classified as ‘’mature and benign’’ or ‘’immature and malignant’’ (referred to as teratocarcinomas)[28,29]. Teratoma formation has been previously reported upon the transplantation of mouse ESC (mESC) derived cardiomyocytes[30], neurons[31] and β-islet cells[32] into immunosuppressed mice, but the initiation and pathogenesis of teratomas are not fully understood[29]. Furthermore, despite the potential elimination of immunogenic concerns associated with allogeneic stem cell sources, additional safety concerns surround the clinical use of iPSCs due to the use of integrating viruses for gene delivery, and the oncogenic properties of the reprogramming factors[33].

Overall, the huge economic burden and high mortality rates imposed by such conditions as heart failure and degenerative diseases requires a new era of therapeutic strategy,

of which regenerative cell therapies show much promise. However, despite global research effort and much advancement in the field there remains a distinct lack of FDA-approved treatments. To date, only one approved stem-cell based product for the treatment of hematopoietic disorders exists (https://www.fda.gov). To prevent this field becoming no more than a pipeline dream for routine treatment a greater understanding of the basic mechanisms of embryogenesis is required. Specifically addressing the molecular pathways governing how ESCs become specialised will inform strategies to efficiently generate homogenous functional cell populations, which upon transplantation will successfully engraft and survive in the host.

1.2 *Early Embryogenesis: formation of the three germ layers*

1.2.1 Fertilisation to blastocyst implantation

The process of mammalian embryogenesis begins from the fertilised egg, which must subsequently proliferate, differentiate to form multiple cell types and generate polarity to enable formation of the body axis[34]. An overview of early embryogenic events describing the processes of how cells undergo fate decisions are given below.

Upon fertilisation the resulting zygote undergoes cleavage events to firstly form 2 identical cells (or blastomeres), which after 3 rounds of this division marks the start of compaction. Compaction is the process whereby bastomeres become attached to each other, interconnected by gap junctions[35], forming a tightly packed spherical structure named the morula (see Figure 1:2). Here, each cell acquires apical-basal polarity and as a result the next cell division can be symmetric (the outside positioned cells give rise to additional outside positioned cells), or asymmetric (the outside cells divide along an in/out axis giving rise to one outside and one inside cell)[36]. Thus, the defining difference between these 2 cell types is that the outside cells has its apical surface exposed, whereas the inside cells has cell-cell contact on all sides. Following 2 major rounds (or waves) of asymmetric cell division (from 8 to 16 cells and 16 to 32 cells) outside cells differentiate to form the trophectoderm (TE) (the precursor to placenta formation)[34]. This arises in part through differential activation of Hippo signalling, a regulator of cell contact-mediated inhibition of proliferation[37], which is activated only in inner positioned cells (surrounded by other cells)[38]. In outside cells Hippo signalling is down regulated, preventing the phosphorylation of the transcriptional co-activator Yes-associated protein, enabling its translocation to the nucleus where it binds TEA domain family transcription factor 4, leading to activation of TE-specific genes such as caudal-related homeobox 2 (Cdx2)[34]. The inner cells form the pluripotent inner cell mass (ICM) and are

associated with the markers Oct4[39], Sox2[40] and Nanog[41]. The expression of TFs are restricted to their respective cells, such that the TE (also marked with Eomes[42] and Gata3[43]) and ICM are incompatible (Cdx2 and Oct4 inhibit each other[44]). Here, the morula transitions to form the early blastocyst, which sees the TE form the fluid-filled blastocyst cavity in a process called cavitation (see Figure 1:2)[45]. This process is associated with epigenetic changes and will become the focus of later sections (see 1.4 and 1.5).

The second cell fate decision sees the cells of the ICM segregated into the extraembryonic primitive endoderm (PrE) (cells become positioned to be in contact with the blastocyst cavity[46]) and the epiblast (Epi)[34]. While Epi cells retain expression of pluripotency genes, the PrE activates the Gata4[47], Gata6[48], Sox17[49] and Sox7[50] TFs. The separation of these lineages is regulated by fibroblast growth factor (Fgf) signalling. Fgf4 is secreted from Epi progenitor cells and interacts with the Fgf receptor 2 on PrE progenitors to activate mitogen-activated protein kinase (MAPK) pathway to induce Gata6 expression[51,52].

At this stage the blastocyst has migrated towards the uterus where it hatches out from the zona pellucida, a glycoprotein rich structure that originally surrounded the oocyte, invades maternal tissues and implants in the uterine wall[53]. Once hatched, the blastocyst adheres to the luminal epithelium, facilitated by integrins and cadherins[54,55], where it invades the stroma at the antimesometrial site of the uterus, proliferates, and subsequently transforms into the egg cylinder[53]. The surrounding stroma cells differentiate to decidual cells (forming the endometrium layer)[56] enabling nutrients and gas exchange and ensuring fetomaternal immune tolerance[53,57]. The TE then proliferates and differentiates into the extraembryonic ectoderm (ExE) to form the proximal half of the egg cylinder and subsequently the placenta[53,58] (See Figure 1:2). Note, the PrE gives rise to the visceral endoderm (VE) which covers both the Epi and ExE compartments elongating the egg cylinder[53] (See Figure 1:2). This provides a foundation for germ layer formation and thus body axis from Epi cells[53].

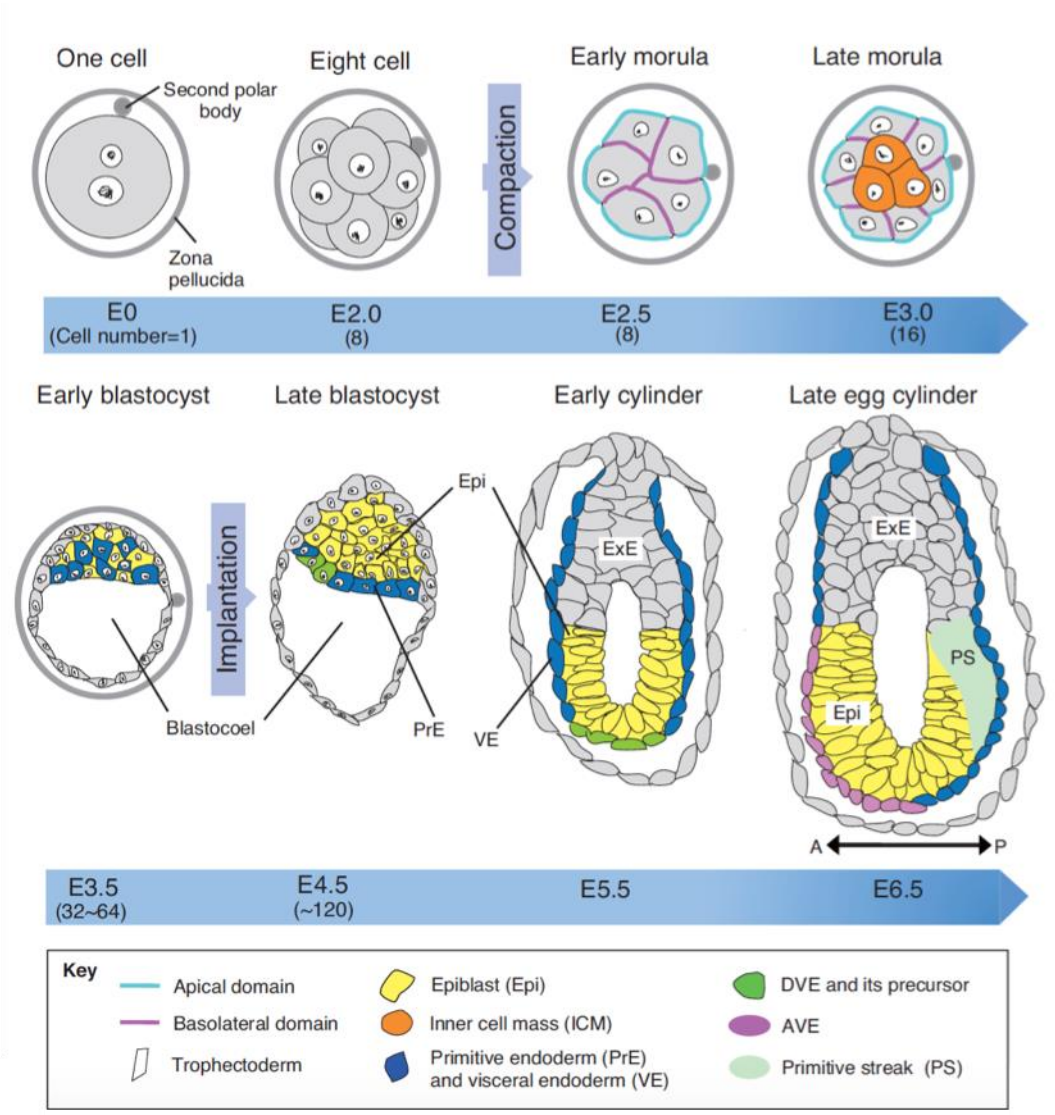


Figure 1:2 Murine Embryogenesis: Fertilisation to Implantation. After 3 rounds of symmetrical cellular division the fertilised egg reaches the 8-cell developmental stage, the point of compaction. Here cells gain apical-basal polarity where they can divide asymmetrically over 2 waves of division to form cell populations of the trophectoderm (TE) and the epiblast (Epi), marking the formation of the early blastocyst. Upon implantation to the uterine wall, the primitive endoderm (PrE) segregates from the Epi and the blastocyst then transitions into the egg cylinder, facilitated by TE differentiation into the extraembryonic ectoderm (ExE). The start of gastrulation is marked by the formation of the primitive streak (PS) from the posterior side of the Epi.

ExE: extraembryonic ectoderm, AVE: anterior visceral endoderm, DVE: distal visceral endoderm. 'A' and 'P' refer to anterior and posterior. Image taken from Takaoka, K and Hamada, H. Development. 2012; 139(1):3-14[34].

1.2.2 Gastrulation

The next major phase of development is gastrulation, a morphogenetic process resulting in the formation of the 3 primary germ layers: ectoderm (outer layer), mesoderm (middle layer) and endoderm (inner layer). These layers form the basis of all tissues types of the body, the ectoderm derives cells of the central nervous system, eye and the skin; the mesoderm forms cells of the heart, blood and kidney; and the endoderm produces cells of the lung, liver and digestive system[59].

In the mouse the start of gastrulation is marked by the formation of the primitive streak (PS) at the posterior side of the Epi. The PS forms from an epithelial to mesenchymal transition (EMT)[60] and provides a gateway for emboly (internalisation) of mesodermal and endodermal cells *via* ingression[61]. The maintenance signals required for PS formation come from reciprocal interactions between the Epi and ExE *via* Nodal and Bone morphogenetic protein 4 signalling respectively[62], thereby ensuring proper elongation of the PS to form the tip or primitive node[63]. Prospective mesodermal cells invade the space between the Epi and the VE at the posterior side of the PS where they become distributed to 4 compartments: axial mesoderm, paraxial mesoderm, lateral plate mesoderm and extraembryonic mesoderm[62]. In comparison, the definitive endoderm arises from the anterior distal region of the PS, in proximity to the node, and intercalates with the VE[64]. It can be noted that from clonal analysis Epi cells are not restricted in their lineage prior to ingression[65], suggesting it likely represents a transient cell population, named the mesoendoderm (which has the potential to form either the mesoderm or the definitive endoderm)[62]. Although such a population has not been found in the mouse embryo, evidence for this bipotential cell has been shown in ESCs[66]. The cells remaining in the Epi, which do not give rise to the PS (found at the anterior side) form the ectoderm.

Gastrulation is marked with specific gene expression patterns that subsequently directs further development of germ layers. For example, Brachyury and Mixl1 are expressed throughout the PS; Foxa2 and Goosecoid are restricted to anterior regions; and HoxB1 and Evx1 are found at posterior locations[67,68]. However, the mechanisms regulating the precise control of spacio-temporal gene expression patterns that results in segregation of cell fates, and thus asymmetry within the developing embryo, is poorly understood.

1.3 *O_2 and embryogenesis*

1.3.1 O_2 availability in embryogenesis

Ambient air has an O_2 concentration of ~21%, which equates to an O_2 exposure for embryonic and adult cells in the range of 2-9%, termed physiological normoxia[69]. More specifically, within the uterine environment, concentrations are typically lower than this and have been reported between 1-5% O_2[70,71]. Hence prior to formation of the uteroplacental circulation (occurring at the end of the first trimester of pregnancy)[72], it can be conceived that embryogenesis transpires at a low [O_2][70]. This low O_2 tension was confirmed in the developing mouse embryo with 2-nitromidazole compounds, which bind to protein and DNA at ≤ 2% O_2[73,74,75,76]. Therefore, delivery of O_2 to the embryo at this stage is mediated through diffusion alone. Owing to the short diffusion distance of O_2, approximately 150 μm from a blood vessel[77,78,79], as demonstrated through an ability to halt tumour development beyond a diameter of ~2 mm (without vascularisation)[80], gradients of O_2 across the developing embryo are likely to form[81]. Thus, cells towards the centre of the developing embryo are exposed to a lower [O_2] than those in outer positions, which are closer to the O_2 source. Intriguingly, even after perfusion of maternal blood from spiral arteries, fetal pO_2 is approximately 50% of maternal levels[70]. This is consistent with detection of isolated low O_2 regions within the mouse embryo, including the developing heart, lung and midbrain, likely attributed to the lack of developing vasculature to meet the rapid growth rate[76].

The concept of O_2 gradients in embryogenesis may in part underlie the basis of morphogenesis: a system of signalling molecules that emanate and diffuse away from the source to form a gradient[82,83]. The idea of a morphogen in development emerged from a collection of studies dating back to the early 20th century[84], and may go some way as to explain asymmetry in the developing embryo. In this regard, the concentration of a morphogenic signal at an exact position within the embryo may function to induce a transcriptional response to determine cell fate. Examples of morphogenic signals are apparent in vertebrate development, for instance: Hedgehog[85], Wingless[86], Decapentaplegic[87], Squint[88] and Fgf[83]. Therefore, it may seem plausible to consider O_2 as a developmental morphogen[69]. However, the precise mechanisms behind how O_2 can function to influence cellular differentiation remains to be fully elucidated.

1.3.2 O_2 sensing within the developing embryo

Adaptive mechanisms have evolved to maintain O_2 homeostasis in multicellular organisms and therefore minimise the potential deleterious effects of O_2 deficiency. Initial or acute responses aim to maximise O_2 delivery to the most critical organs such as the heart or brain *via* increasing cardiac output, hyperventilation and arterial vasodilation[89]. In comparison, chronic adaptive responses refer to alterations in gene expression for instance activation of glucose metabolism, erythropoiesis, angiogenesis[89]. To facilitate such adaptive responses, changes in [O_2] must be sensed. Candidates for potential O_2 sensors include NADPH oxidases[90], the mitochondria[91] and specific ion channels[92,93]. However, central to the role of O_2 sensing are the hypoxic inducible factors (HIFs), and evidence for their role in developmental processes are discussed below.

HIFs are heterodimeric TFs consisting of a constitutively expressed HIF-1β subunit and an O_2-sensing HIF-α subunit (HIF-1-3α)[94]. In environments of an [O_2] greater than 5%[95], the HIF-α protein becomes hydroxylated at 2 proline residues by a prolyl hydroxylase domain (PHD)[96]. The 3 isoforms of PHDs (PHD1, 2, 3) belong to the 2-oxoglutarate dependent dioxygenase (2-OGDO) family (discussed below) and crucially require molecular O_2 as a cofactor[97]. Hydroxylation acts as a recognition signal for the binding of the von Hippel-Lindau (pVHL) ubiquitin E3 ligase complex, which functions to induce HIF-α polyubiquitination resulting in subsequent proteosomal degradation[94]. Furthermore an additional hydroxylation event occurs at asparagine residues by factor inhibiting HIF (FIH) (also a member of the 2-OGDO family), which prevents the binding of the HIF-α transcriptional coactivator CBP/p300 thereby repressing HIF-α transcriptional activity[98]. Upon a reduction in [O_2] the rate of prolyl and asparaginyl hydroxylation is reduced, consequently stabilising HIF-α and thus enabling its translocation to the nucleus and subsequent dimerisation with HIF-1β subunit[96]. This HIF-α/β dimer binds to hypoxic response elements (HRE), characterised by the core DNA motif G/ACGTG, to induce gene transcription[96]. Over 60 HIF target genes have been identified, and were found to be involved in processes including, angiogenesis, vasodilation, and glycolytic metabolism activation[99,100].

The impact of HIF-signalling upon development can be best understood from genetic ablation studies. Loss of the HIF-1β *in vivo* results in embryonic lethality by embryonic day (E) 10.5, attributed to defective angiogenesis[99,100]. Similarly, silencing of the HIF-α subunit *in vivo* is also lethal by E 11, manifested by neural tube defects and cardiovascular malformations[101,102]. In addition, silencing of Cited2, a negative regulator of HIF-1α *via* competitive binding to CBP/p300, also has lethal effects through

enhancement of HIF response genes[103]. Collectively, these studies demonstrate that HIF-mediated O_2-sensing plays a critical role in embryogenesis.

Specifically, a fundamental role of HIF is in the development of the placenta[76], which forms in the mouse at E 10.5-11.5 and is required for O_2 and nutrient delivery to sustain embryo growth. Placental defects arising from the loss of HIF include reduced chorioallantoic interaction (an interaction between the embryo and forming placenta)[104], limited vascularisation of the placenta[105] and reduced development of the labyrinthine layer (site at which fetal blood vessels and maternal blood sinuses are bought in proximity)[106]. As embryogenesis progresses the Exe expands to form the chorion and ectoplacental cone, which contain trophoblast progenitor cells. Trophoblast cells proliferation and differentiate into distinct cell subtypes to support placental development. HIF activity, demonstrated through ablation of HIF-1β, HIF-1α or HIF-2α, was shown to be required to drive spongiotrophoblast specification (required for structural support to the labyrinthine) through regulation of the Mash2 TF[104]. Therefore, this not only suggests there are potential overlapping roles of the HIF family members, but also directly demonstrated that HIF can regulate cell fate decisions. Related to this, [O_2] has shown to be a direct regulator of placental development. An [O_2] mimicking the uterine environment before and after connections are developed with the maternal vasculature (2.5% vs 8.6% O_2)[69] was shown to control cytotrophoblast (single cell inner later for the trophoblast) differentiation into an invasive phenotype, which is required to aid the establishment of maternal-fetal circulation at the placental bed[107]. Importantly, this highlights the impact of O_2 gradients on cellular differentiation.

An additional system in which HIF is intrinsic to is the formation of the cardiovascular system. Silencing of HIF-1β resulted in decreased vascular endothelial growth factor (VEGF) messenger RNA (mRNA) and protein expression, which resided in vascular defects[99,100]. Furthermore, HIF-1α-deficient embryos showed arrested cardiac morphogenesis, recorded at various stages of development[108,109]. Unlike HIF-1α or HIF-1β, HIF-2α deficient embryos survive until mid-to-late gestation, and have independently been shown to present with catecholamine dysregulation[110], cardiac hypertrophy[111] and a lung maturation defect[112]. These experiments therefore demonstrate that HIF-1 and HIF-2 perform separate developmental functions.

HIF-signalling, exclusively *via* HIF-2α, has been shown to regulate stem cell function by regulation of Oct4 and additional pluripotency markers such as Sox2 and Nanog[113,114,115]. In support of this notion, HIF-2α-silenced embryos showed a loss of primordial germ cells (PGCs), which require Oct4 for survival[113]. In addition Cited2 has been shown to

control expression of the pluripotency genes Nanog, Tbx3, Klf4 and interesting its ablation induced spontaneous ESC differentiation, suggestive of a role in stem cell maintenance[116]. In comparison, HIF-1α has been implicated in the maintenance of neural stem cells *in vivo*[117] and is linked to the commitment of ESCs to the neural lineage *via* direct activation of Sox1[118]. In agreement with the *in vivo* role for HIF-1α in cardiogenesis as described above, exogenous expression of HIF-1α in mESCs promoted cardiac differentiation, demonstrated by lineage marker specific expression and an enhanced beating phenotype of EBs[119]. Overall, this suggests O_2, and hence its downstream signalling pathways, are required for the regulation of stem cell function. Note, HIF-3α is the least studied of the HIF protein subunits and is subject to extensive splicing, displaying different O_2 sensitivities and abilities to dimerise HIF-1β[120]. The precise biological functions of HIF-3α are unknown, but it has previously been proposed in a hESC study to regulate the long-term response to low [O_2], after the initial transient response from HIF-1α has elapsed[114].

1.3.3 2-oxoglutarate dependent dioxygenases

The 2-OGDO domain containing proteins (PHDs and FIH) are responsible for the O_2 sensing capabilities of HIFs, as mentioned above, and are highly conserved within many other proteins. The 2-OGDO family are comprised of over 60 members in humans alone[121], and can catalyse oxidative transformations in a diverse array of biological processes within mammalian cells including: post-translational modification of collagen, fatty acid metabolism and epigenetic regulation[122].

2-OGDO's are soluble, non-heme, Fe^{2+} containing proteins[123]. Activation of these enzymes begins with binding of 2-oxoglutarate (2-OG) (also known as α-ketoglutarate) and Fe^{2+} in the catalytic domain[124]. Next, O_2 binds to Fe^{2+} to induce oxidative decarboxylation of 2-OG to succinate and CO_2, which results in the formation of an oxidised ferryl intermediate ($Fe^{3+/4+}$)[124] (illustrated in Figure 1:3). During this catalysis an O atom becomes incorporated into the alcohol group (-OH) of the oxidised prime substrate and into succinate[96]. Enzymatic activity is restored by recycling $Fe^{3+/4+}$ back to Fe^{2+} by the reducing agents ascorbic acid (AA)[125] and glutathione (GSH)[126]. Fe^{2+} is ligated by 2 histidine residues and a carboxylate (named the facial triad[96]) at the catalytic site, which is positioned within a distorted double stranded β-helix (DSBH) fold consisting of 8 strands (referred to as a jelly roll) that supports the binding of Fe^{2+} and 2-OG[122]. Owing to the varying number of residues between the 4th and 5th strands of the β-helix, the jelly roll structure can distinguish 2-OG subfamilies. FIH has a long insert (thought to

be involved in substrate recognition) and is part of the Jumonji C (JmjC) family of histone modifying enzymes, which have a characteristic squashed DSBH barrel structure (JmjC fold)[127,128]. PHD2 has no insert and comes from the same structural subfamily as the collagen prolyl hydroxylases (the first identified pathway involving 2-OGDO), which have a cupin barrel formation[127,128].

Due to the co-substrate requirements to sustain the activity of 2-OGDOs, these enzymes are considered cellular sensors for changes in metabolites, O_2 availability and Fe^{2+} redox status[124]. Given the low O_2 tensions present during early embryogenesis and the large number of enzymes that share the 2-OGDO domain, it is reasonable to believe that aside from HIF regulation other proteins may be involved in developmental events. One emerging novel area is the role 2-OGDOs have in the regulation of epigenetic function in response to changes in [O_2], which is addressed in the sections below.

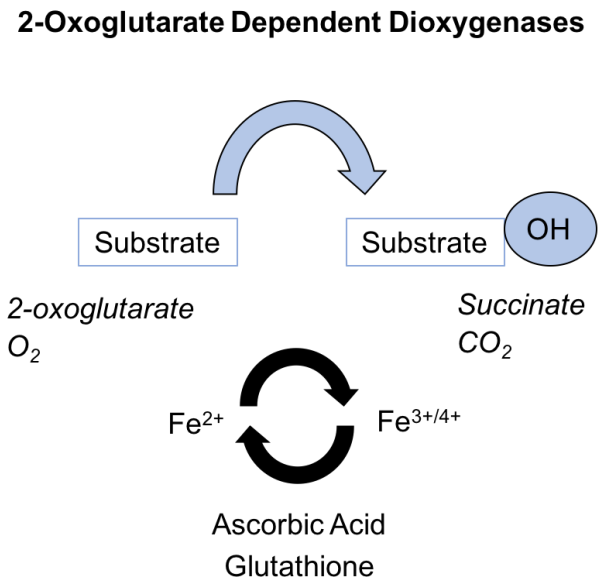


Figure 1:3 The catalytic 2-oxoglutarate dependent dioxygenase reaction. 2-oxoglutarate and O_2 are required as co-substrates for the hydroxylation of target molecules by 2-oxoglutarate dependent dioxygenase domain containing proteins. This reaction yields succinate and CO_2 while also generating oxidised ferryl intermediates. $Fe^{2+/3+}$ is recycled back to Fe^{2+} by ascorbic acid (AA) and glutathione (GSH) to restore enzymatic activity.

1.3.4 The effects of O_2 on embryonic stem cell function

Historically, cell culture was focussed upon the exact requirement of nutrients and growth factors to maintain cell growth. However, in an *in vivo* environment, cells are exposed to a much lower partial pressure (pO_2) than that found in a humidified cell incubator, which consists of an atmosphere of 95% air/5% CO_2 (equating to ~20% O_2). Therefore, to better resemble more physiological conditions such as those representing the stem niche, manipulation of O_2 concentrations has been the focus of much research.

Measurements of O_2 tensions in cellular compartments have defined physiological normoxia in a range of 2-9% O_2[69]. However, concentrations have been recorded to be as low as 1% or less among stem cell niches[81]. Exposure of hESCs to 21% O_2 (ambient air) promoted spontaneous differentiation, whereas if maintained under low O_2 conditions (3% or 1% O_2) cells maintained pluripotency[129]. A similar study demonstrated culture of hESCs in 5% O_2 (in contrast to 21% O_2) acted to maintain long-term self-renewal of cells through elevated Nanog expression[130]. In addition, culture of hESCs in 5% O_2 was shown to prevent X chromosome inactivation, the process of transcriptionally silencing one of the X chromosome copies in female mammals, enabling ESCs to be maintain in a developmentally more immature state[131]. Further to enhanced chromosomal stability and clonality when cultured under physiological normoxic [O_2], microarray profiling revealed hESCs to have minimal activation of developmental pathways when cultured in 2% O_2, compared to ambient air[132]. However, these findings are controversial as it has been shown in separate studies in hESCs that exposure to 5% O_2[133] and 4% O_2[134] had no effect on stemness marker expression, compared to culture in 21% O_2.

Furthermore, culture of mouse preimplantation embryos in 5% O_2 demonstrated fewer gene expression perturbations, compared to control embryos exposed to atmospheric O_2[135]. This O_2 tension has independently been shown to enhance embryo development, suggesting low O_2 tensions may enhance proliferation[136]. Despite this, 1% O_2 has been shown to reduce self-renewal and induce differentiation of mESCs *via* downregulation of the LIF-Stat pathway *via* HIF-1α[137]. Spontaneous differentiation of mESCs towards the endodermal lineage was observed in 1% O_2 (but not in 3% and atmospheric O_2 conditions)[138], which was in agreement with the effect of 5% O_2 under a directed differentiation protocol[139].

Discrepancies in the literature over the exact function of O_2 in pluripotency regulation are apparent, likely due to the source of ESCs, culture conditions and the time exposed to

low [O_2]. However, collectively it is accepted that a more 'hypoxic environment' promotes the undifferentiated state in stem and precursor cell populations, but the mechanisms underpinning this process are still largely not understood[69]. Therefore, O_2 is likely to be an intrinsic signalling molecule in early differentiation events.

1.4 *Epigenetics*

The process of embryogenesis requires ESCs to lose pluripotency and differentiate towards specific cell lineages. During this process transcriptional programmes, which can be controlled through epigenetic modifications, define cellular identity[140]. The term epigenetics refers to heritable and reversible changes to gene expression (and hence cell function), which do not involve alterations to the DNA sequence[141]. Hence, despite most cells within multicellular organisms sharing an identical genotype, epigenetics enables the generation of diverse array of cell types with distinct functions during development[141]. Epigenetics can control gene transcription through regulation of chromatin compaction, which consequently controls the levels of TF accessibility to the DNA. The cellular 'epigenetic landscape' or signature is inherited across cell divisions, imparting an epigenetic memory, functioning to maintain the cellular transcriptional profile and thus specificity[142].

1.4.1 Epigenetic modifications

In the nucleus 147 bps of DNA is wrapped around an octamer of histones, consisting of 2 H3-H4 histone dimers that are surrounded by 2 H2A-H2B dimers, to form the nucleosome, the basic subunit of chromatin[143]. The spacing of nucleosomes, which are separated by linker DNA, determines the chromatin structure[143]. Regions that are tightly packaged are referred to as heterochromatin, while the loosely packaged areas are named euchromatin, and are thus associated with transcriptional silencing and activation respectively[144]. The level of chromatin compaction can be controlled by 4 main types of epigenetic modification: adenosine triphosphate- (ATP) dependent remodelling complexes (ADCRs), non-coding RNAs (ncRNA), histone modifications and DNA methylation[141]. These chromatin remodelling complexes often function in combination with each other to enable precise and diverse transcriptional control to obtain specified cell types[141,145].

ADCRs utilise energy from ATP hydrolysis to destabilise the nucleosome by altering its position, changing the conformation of nucleosomal DNA and the composition of the histone octamer[146]. This nucleosome remodelling can thus lead to transcriptional

activation or repression, depending upon TF accessibility to DNA[147]. These complexes can be characterised into 4 groups based upon their core ATPase subunits, and are essential for aspects of embryonic development (reviewed by Hargreaves, D and Crabtree G (2011)[148]). ncRNAs, which make up greater than 80% of transcribed genes, are also implicated in the control of gene expression[149]. ncRNAs are divided into 2 groups based on size: small ncRNAs (<200 nucleotides) encompassing microRNAs (miRNAs), piwi-interacting RNAs (piRNAs) and short interfering RNAs (siRNAs); and long ncRNAs (lncRNAs) (>200 nucleotides)[150]. ncRNAs regulate gene expression post-transcriptionally and have also been found to be associated with chromatin remodelling complexes to facilitate the alteration of DNA methylation and histone status[151] (reviewed by Wahlesedt, C and Peschansky, V (2014)[152]).

Histones can be post-translationally modified covalently at their protruding tails (Figure 1:4) predominantly by acetylation, phosphorylation and methylation. These dynamically expressed marks can be found in combination on the same histone tail, encompassing what is described as the 'histone code'[153]. Acetylation, controlled *via* 2 opposing enzyme families: histone acetyltransferases (HATs) (addition) and histone deacetylases (HDAC) (removal), occurs on lysine residues and is associated with euchromatin[154]. Phosphorylation of serine, threonine and tyrosine residues also functions to control chromatin compaction, regulated through the action of kinases and phosphatases that add and remove this modification respectively[155]. The methylation of histones occurs at lysine (K) residues on histones H3 and H4, and can be presented in either a mono, di or trimethylated (me2/3) form (arginine residues can also be modified by 1 or 2 methyl groups)[156]. This array of modifications results in a huge potential for the regulation of gene function, and have subsequently been shown to be associated with gene activation and repression. For example, methylation at H3K4, H4K36 or H3K79 is linked with transcriptional activation, whereas at H3K27, H3K9 or H3K20 is associated with transcriptional repression[149]. Histone methylation is also controlled by the function of 2 enzymatic processes; histone methyltransferases (HMTs) and histone demethylases (KDMs), to add and remove methyl marks respectively[149].

Perhaps the most well studied of the epigenetic modifications is DNA methylation. DNA methyltransferase (Dnmts) catalyse the covalent addition of a methyl to the 5th carbon residue of cytosine (Figure 1:4)[157]. These enzymes incorporate the methyl from the donor S-adensylmethionine (SAM), at CG dinucleotides (CpG) within the genome, forming the 5-mC mark and S-adenosylhomocysteine (SAH)[157]. Clusters of CpG dinucleotides are typically found at CpG islands within gene regulatory elements, for example approximately 70% of gene promoters in the human genome were found to

have a high CpG content[158]. *De novo* methylation conducted by Dnmt3A and Dnmt3B (maintenance of methylation performed by Dnmt1)[159] functions to repress gene transcription through blocking recognition sites for TF binding[143]. Alternatively, the 5-mC mark can recruit methyl-CpG-binding domain (MBD) proteins, which subsequently function to assign histone deacetylases to facilitate gene repression[160].

Unlike the previously stated epigenetic modifications, the ability to adjust levels of the 5-mC through a demethylation pathway is less well characterised. Formerly, the removal of the 5-mC mark was believed to occur passively during replication in the absence of Dnmt1[161]. However, due to evidence for rapidly changing methylation events occurring during embryonic development (mentioned above) and in somatic cells, an active demethylation process was proposed[161]. Multiple candidate mechanisms have been proposed: cytidine deaminases[162], DNA glycosylase-mediated base excision repair (BER) pathway[163], nucleotide excision DNA repair[164] and even Dnmt enzymes themselves[165]. However, the lack of a unifying active DNA demethylation pathway questions the significance of these mechanisms, suggesting each pathway may be specific to an individual biological system[166]. But, recent advancements in this field have now led to the identification of a novel direct oxidative DNA demethylation pathway.

1.4.2 DNA demethylation: formation of a novel epigenetic mark

In 1972 an additional modified form of cytosine was discovered in mammalian DNA, named 5-hydroxymethylcytosine (5-hmC)[167]. Originally perceived as an oxidative-damage product of DNA, 5-hmC was largely ignored until 2 independent pioneering publications in 2009 demonstrated 5-hmC was readily detected in mESCs and in terminally differentiated neuronal cells[168,169]. Levels of 5-hmC in mammalian tissues have subsequently been reported to be prevalent in typically 0.1% of cytosine (C) residues, with the largest levels of up to 0.7% recorded in brain[169,170].

The 5-hmC mark is generated from oxidation of the 5-mC *via* ten-eleven translocation (Tet) enzymes (Figure 1:4). Tet1, the founding member of this enzymatic family, was initially discovered in acute myeloid leukaemia as a fusion partner of the histone methyltransferase (MLL) gene[171]. However, it was the pioneering work by Rao, A *et al* (2009)[168,172] that identified the potential 5-mC modification capabilities of Tet proteins, due to their identification as homologs to the trypanosome J-binding proteins, which function to oxidise thymine to hydroxymethyluracil[166]. Currently, this Tet enzymatic family consists of 3 members (Tet1, Tet2 and Tet3), which all share a core catalytic domain at their C-terminus (Figure 1:5A)[173]. This domain features a cysteine-rich insert and a larger

DSBH fold, which exhibits 2-OGDO activity, and thus for catalytic function has an absolute requirement for molecular O_2 (Figure 1:5B)[173]. In addition, at the N-terminus, Tet1 and Tet3 contain a CXXC zinc-finger binding domain (Figure 1:5A)[166]. The precise function of this domain is elusive, although has been suggested to facilitate binding to CpG regions and thus targeting to specific gene loci[174,175,176]. Tet2 lost its CXXC domain during evolution due to a chromosomal inversion[172]. However, this separated region encodes for the CXXC-containing gene product Idax, which binds to CpGs and interacts with the catalytic domain of Tet2[177].

The Tet-mediated erasure of the 5-mC mark occurs as a result of 2-OGDO activity to yield 5-hmC. Investigations into the stability of 5-hmC suggested this mark is not a transient intermediate of demethylation, and in fact may itself regulate chromatin structure and transcription[178,179], leading to suggestions that it may be considered the 6th base of the genome (whereby the 5th base is 5-mC)[180]. 5-hmC has been shown to be positioned within gene regulatory regions[181] and in some cases has been shown to positively correlate with gene transcription[182,183,184]. This may be due to remodelling of the chromatin structure into a euchromatic form (Figure 1:4)[185,186], potentially through influencing the thermodynamic properties of the DNA helix[187]. Alternatively, 5-hmC has been shown to prevent the recruitment of MBD proteins and thereby relieving their transcriptional repressive effects[188]. Tet proteins may also induce transcriptional activation by facilitating the generation of the H3K4me3 mark *via* a physical association with O-GlcNAc transferase (Ogt)[189,190]. Ogt is required for the activation of Host cell factor 1 (Hcf1), a component of the Set protein methyltransferase complex[191,192,193] (see further details in Chapter 4: Results 2 Introduction). Nonetheless, Tet proteins have also been reported to have potential transcriptional repressive effects. Genetic ablation of Tet1 in mESCs was found to upregulate a significant fraction of genes that are normally silenced by Polycomb Repressor 2 (Prc2) mediated H3K27me3 generation[194] (see further details in Chapter 4: Results 2 Introduction). Furthermore, transcriptional repression may be achieved through association of Tet1 with the Sin3A corepressor complex to mediate histone deacetylation[195,196]. Collectively, this fundamentally highlights the multiple transcriptional regulatory roles surrounding this novel Tet/5-hmC pathway.

To restore the unmodified C base 5-hmC can be removed in a replication-dependent manner, or subjected to further successive oxidation reactions by Tet enzymes to form 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) (Figure 1:5A)[166,197,198,199]. These intermediates can be subjected to active removal by a DNA repair pathway, specifically involving the use of the DNA repair protein thymine DNA glycosylase (TDG)[166]. TDG recognises and excises 5-fC and 5-caC (but not 5-mC or 5-hmC) at the glycosidic bond,

generating an abasic site that is subsequently repaired by BER[200] to restore C[166]. This mechanism is consistent with the low abundance of these intermediates in the genome, whereby it has been reported in mESCs, that 5-hmC is approximately 10-100-fold more prevalent than 5-fC and 5-caC[201]. In line with this, genetic ablation of TDG in mESCs resulted in a 5-10-fold elevation in both 5-fC and 5-caC marks[202,203,204,204,205], suggesting that these are short-lived intermediate epigenetic marks. Furthermore, despite all oxidative substrates having similar Tet binding affinities, Tet enzymes display different substrate preferences, such that Tet enzymes are less active on 5-hmC and 5-fC substrates[198,206]. This may be attributable to the adoption of a more restrained conformation in the active site, which prevents hydrogen bond abstraction and thus reduces catalytic efficiency[206]. This therefore further suggesting 5-hmC is the most stable and abundant oxidative mark. Alternatively, the presence of oxidative marks can prevent *de novo* methylation. Following DNA replication, hemi-methylated CpGs are transiently formed, with only the parental strand containing the 5-mC mark[207]. Recruitment of Dnmt1 is required to restore original methylation patterns in the newly synthesised DNA strands[207]. The presence of oxidative marks prevents the ability of Dnmt1 to methylate CpGs, resulting in progressive dilution of 5-mC over successive rounds of DNA replication[208,209,210].

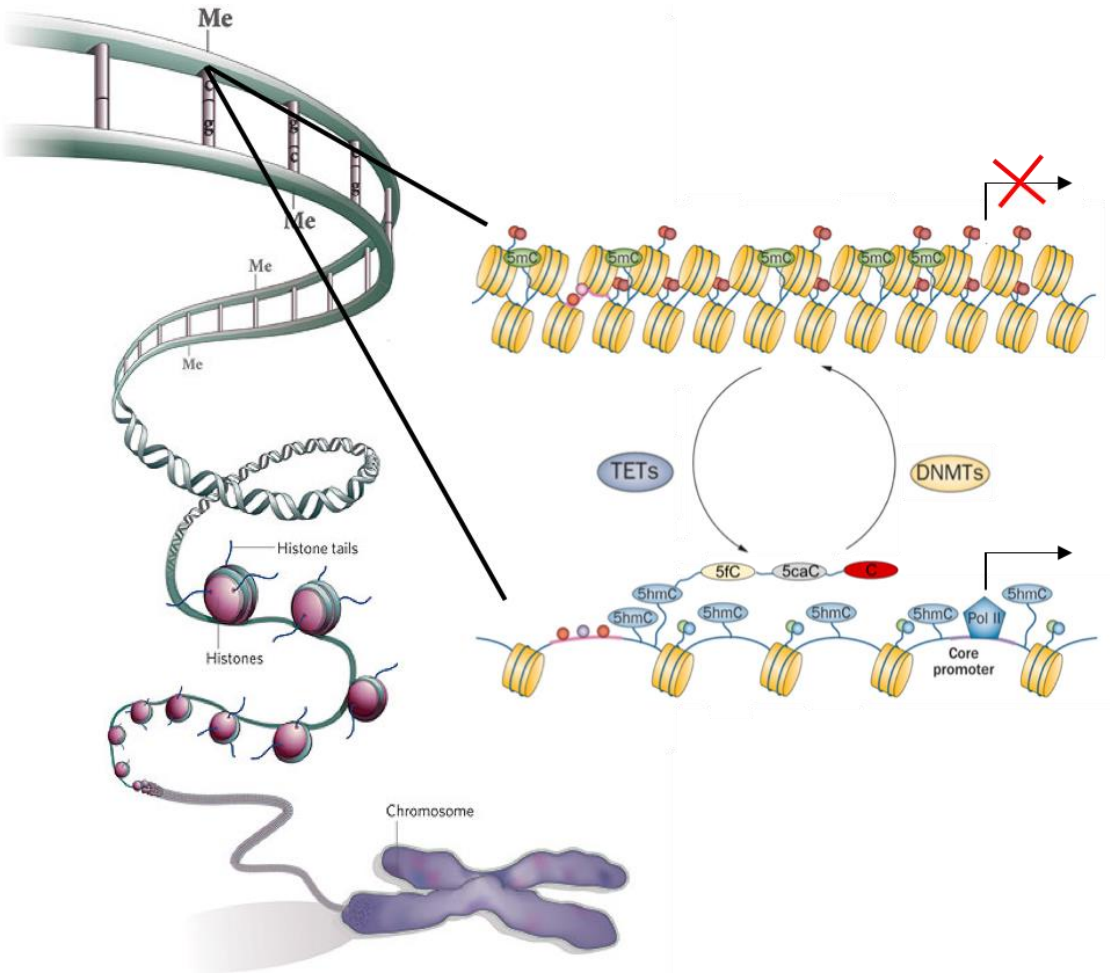


Figure 1:4 Chromatin and DNA methylation. The chromosome is made up of DNA packaged in chromatin, which is formed from nucleosomes, separated by linker DNA. The level of chromatin compaction, affecting transcription factor accessibility, can be dynamically adjusted through epigenetic modifications. For example, histone tails can be modified by acetylation, methylation and phosphorylation, or the DNA itself can undergo methylation. Specifically, the addition of a methyl group to the 5th carbon residue of cytosine (C), *via* DNA methyltransferases (Dnmts), is associated with a condensed (or heterochromatic) chromatin structure, and is representative of transcriptional repression. Demethylation (oxidation) of the methylated cytosine mark (5-mC) by ten eleven translocation (Tet) enzymes form 5-hydroxymethylcytosine (5-hmC), which relaxes the chromatin structure into its euchromatic form, permitting gene transcription. The 5-hmC mark can undergo successive oxidation reactions *via* Tet enzymes, to form 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC). These oxidative intermediates can subsequently undergo DNA repair to regenerate unmodified C. Note the coloured circles represent histone marks, which can also influence chromatin structure, see main text for specific details. Adapted from: Rajender, S *et al*. Mutat Res. 2011; 727:62-71[144] and Greco, CM and Condorelli, G. Nat Rev Cardiol. 2015; 12:488-497[149].

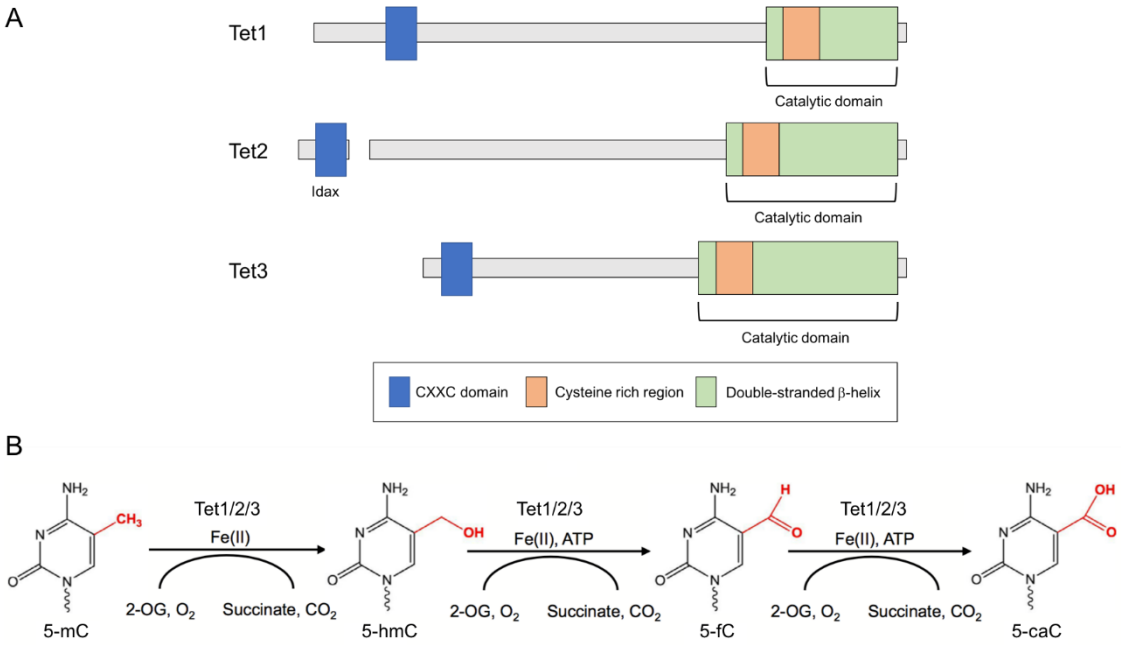


Figure 1:5 The structure and mechanism of Tet enzymes. (A) Schematic structural representation of the 3 ten eleven translocation (Tet) enzymes. All enzymes contain a C-terminal catalytic domain containing a cysteine rich region and a double stranded β-helix harbouring 2-oxoglutarate dependent dioxygenase activity. Tet1 and Tet3 additionally contain a CXXC DNA binding domain at their N-terminal region. Due to a chromosomal inversion event during evolution, Tet2 lost its CXXC domain. This separated region encodes for the CXXC-containing protein Idax and can functionally associate with Tet2. (B) Tet1/2/3 mediate the successive oxidation of methylated cytosine (5-mC) into 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC). This catalytic reaction, *via* the 2-oxoglutarate dependent dioxygenase domain, requires 2-oxoglutarate (2-OG) and O_2, which in the presence of Fe^{2+} are used to oxidise the substrate, yielding succinate and CO_2. Note ATP is required for further oxidation of 5-hmC. (A) Adapted and (B) taken from Tan, S and Shi, Y. Development. 2012; 139:1895-1902[173].

1.5 *Epigenetics in early mammalian development*

A fundamental question of developmental biology surrounds the signalling mechanisms that instruct cell fate decisions during embryogenesis. Such cell specification events have been linked to chromatin regulation, resulting in a hypothesis that epigenetic asymmetry functions as a driver for cell lineage allocation[211].

Perhaps one of the most significant developmental epigenetic events occurs in the process of X-inactivation. Both X-chromosomes express Xist, a large untranslated RNA, which coats the X-chromosome from which it is expressed and triggers genetic silencing[212]. DNA methylation silences one of the Xist genes[213], and it is this chromosome from which genes will be expressed in future progeny cells. Changes in global methylation levels also occur during early embryogenesis. After fertilisation a global loss in DNA methylation occurs in the zygote[214]. Methylation patterns become re-established *de novo* in the developing embryo in all but CpG islands[214]. Hypomethylation of these CpG islands remains until later developmental stages when cellular specification arises. Here, methylation is acquired and functions to silence germline-specific genes (including pluripotency genes), correlating with induction of differentiation[215,216]. Furthermore, developmental genes Pax and Hox, have reported to show cell type specific DNA methylation at CpGs[217], leading to the hypothesis that differential methylation may serve to regulate gene expression during differentiation[218]. In fact, throughout development, the oscillating levels of methylation have recently been recognised to contribute to the left-right asymmetric body plan during vertebrate embryogenesis[219]. A second wave of DNA demethylation occurs during germ line formation. PGCs, which arise from the extraembryonic mesoderm cells of the Epi, undergo a reprogramming event for erasure of genomic imprints for formation of mature sperm and oocytes[220,221]. The importance of DNA methylation in development was made apparent through genetic ablation of Dnmts in mice, whereby Dnmt1 silencing resulted in embryonic lethality at E 8.5-9[222]. In comparison, Dnmt3A-null embryos appeared normal at birth, but died at approximately 1 month of age, whereas Dnmt3B ablated embryos resulted in no viable mice at birth, attributed to growth impairment and defects in neural tube formation[223]. Furthermore, in the embryos of patients suffering from early pregnancy loss, Dnmt1 expression and DNA global methylation levels were downregulated, which was associated with abnormal embryo implantation and development[224].

Epigenetic events also control genomic imprinting, the process of enabling a small proportion of genes (150 in mouse, 75 in human[225]) to be expressed in a parent-of-origin-

specific manner (compared to the vast majority of genes, which are expressed from both alleles simultaneously)[221]. Differential DNA methylation between the 2 parental chromosomes is the main mark defining imprinted genes and remain present during global DNA demethylation events after zygote fertilisation[221]. Interestingly, a trimethylated-histone mark signature (H3K4me3, H3K9me3 and H4K20me3) has also been identified in imprinting control regions[226], suggesting potential cross-talk between epigenetic modifications. The presence of H3K4me3 and H3K27me3 marks at specific chromatin regions (bivalent domains) (see Chapter 4: Results 2 Introduction) were also detected on the same allele in imprinted genes[226]. These bivalent domains are also present within the TE and to a greater extent in the ICM of mouse blastocysts of developmental and differentiation-related genes[227]. This is attributable to a greater abundance of H3K27me3 mark in the ICM, which was determined to be a fundamental epigenetic component for inducing the segregation between the ICM and TE, and therefore demonstrates epigenetic asymmetry in the embryo[221,228]. Furthermore, the H3K9me3 mark was identified as a repressive mark for TE development, whereby H3K9me3 expression was inversely proportional to expression of TE-related genes as they undergo differentiation[229]. In support of this evidence for the presence of lineage-specific histone marks during embryogenesis, a gradient of the H3R26me2 mark across the 4-cell blastomere stage was reported, whereby high levels of this mark were found to form the pluripotent cells of the ICM, compared to low expression that formed the TE[228,230].

1.5.1 The functional role of Tets in early mammalian development

To examine the relevance of Tet enzymes as potential O_2 sensors during embryogenesis their ability to have functional roles in developmental processes must be considered.

The importance of Tet enzymes in physiology was first discovered from myeloid malignancies, whereby mutations of Tet2 were reported in acute myeloid leukemia (AML), myeloproliferative neoplasms (MPN) and chronic myleomonocytic leukemia (CMML)[207]. The majority of mutations are inactivating, disrupting catalytic function and thus inducing aberrant DNA methylation patterns[207]. Interestingly, hematopoietic stem/progenitor (HSPC) cells bearing Tet2 mutations from MPN[231], and HSPCs genetically ablated of Tet2, demonstrated a skewed differentiation towards myeloid lineages[232]. These events have also been recapitulated in various mouse models[207], which highlights the influence Tets can have upon cell differentiation. Furthermore, in AML and MPN further inhibition of the Tet-mediated active DNA demethylation process can transpire due to mutations in isocitrate dehydrogenase (IDH) enzymes[233,234]. IDH

enzymes produce 2-OG, a co-substrate of the 2-OGDOs, in the tricarboxylic acid cycle (TCA), however mutations in IDH result in the synthesis of the oncometabolite 2-hydroxyglutarate (2-HG)[235]. 2-HG functions as a competitive inhibitor of Tet enzymes (and JmjC-KDMs) culminating in the loss of 5-hmC levels (and histone methylation)[236]. This loss of 5-hmC is also associated with other cancer types including melanoma, colorectal, gastric and glioblastoma, and is therefore considered a hallmark of cancer[237,238,239].

The first instance of Tet involvement in embryogenesis was reported during the global loss of DNA methylation in the zygote. Here, 5-mC becomes oxidised to 5-hmC predominantly in the paternal genome, while the maternal genome is protected from this oxidation stage *via* an interaction between H3K9me2 and the oocyte derived factor Pgc7[173,240]. Paternal oxidation is catalysed by maternal Tet3, owing to its high expression (prior to cleavage events where expression is lost) relative to Tet1 and Tet2 in the fertilised zygote[241,242]. This increase in 5-hmC was abrogated by silencing of Tet3 in mouse oocytes[243]. Furthermore, Tet3-deficient zygotes showed reduced expression of the Oct4 pluripotency marker during the morula developmental stage, which in some animals resulted in abnormalities in E 11.5 and failed to develop[242]. However, it has since been suggested that Tet3 haplo-insufficiency may be the cause of these aberrant phenotypes induced by the maternal loss of Tet3, rather than defective 5-mC oxidation, suggesting Tet3 is required for neonatal growth[244]. This study further demonstrated that the paternal genome can become hypomethylated independent of maternal Tet3, suggesting existence of compensatory mechanisms[244].

A second developmental transition involving the loss of 5-mC occurs in PGCs, which can be divided into 2 waves. In mouse (but also conserved in human) between days E 7.25 and E 9.5 global levels of 5-mC are lost through passive dilution, but then in addition acquire active removal *via* Tet1-mediated oxidation up to E 13.5[236]. Immunostaining showed 5-hmC was induced after E 9.5, which in combination with the presence of BER machinery is suggestive of Tet involvement to restore unmodified C[245,246]. mRNA expression analysis in PGCs (from E 9.5) revealed that Tet1 was most abundant paralogue[246] and from Tet1-deficient mice it was shown that this hydroxylation is required for imprinting erasure in germline cells (new sex-specific epigenetic modifications are established as development progresses)[247]. It should be noted that global methylation in germ cells is largely unaffected by the loss of Tet1 alone[247] or in combination with silencing of Tet2[248], indicative of Tet-mediated oxidation having a locus specific effect and/or that passive demethylation is accountable for the global 5-mC loss[236].

Staining for 5-hmC in the pre-implantation mouse embryo revealed detectable levels in all stages (2-cell to blastocyst), but maximal signal intensity was reported in the ICM of the blastocyst, which appear matched to Tet1 and Tet2 mRNA expression[173,249]. This detection of 5-hmC within ESCs has been further confirmed separately by thin layer chromatography (TLC)[168] and immunofluorescence[186] techniques. Suggestions of Tet involvement in development was first realised *via in vivo* teratoma formation assays, a method of determining cell pluripotency *via* assessing the ability of cells, upon injection into immunosuppressed mice, to form tumours (teratomas) comprising tissues representative of all 3 germ layers[250]. ESCs ablated of Tet1, and in combination with Tet2 loss[248], formed large haemorrhagic teratomas containing trophoblast cells[251] and increased endodermal and loss of neuroectodermal differentiation makers[252]. Expression analysis confirmed the induction of TE markers Cdx2, Eomes and Elf5, indicative of Tet protein involvement in supressing TE formation[252]. However, *in vivo* the functional loss of either Tet1 or Tet2 in mice is compatible with development and animals appear healthy throughout adulthood, suggesting deficiency in either Tet does not compromise developmental potential[251,253 232]. Combined loss of Tet1 and Tet2 negates 5-hmC (to greater extent than when either Tet alone is silenced)[252], which despite retaining pluripotency in ESCs does reside in a fraction of mice displaying midgestation abnormalities and perinatal lethality[248].

More recent studies sought to elucidate the developmental potential of mESCs displaying complete loss of Tet activity. Triple Tet silenced (TKO) mESCs were completely depleted of 5-hmC and showed restricted developmental potential, attributed by a lack of endodermal and mesodermal markers as well as immature ectodermal structures[254]. The differentiation potential was clarified *in vivo*, whereby GFP targeted TKO and Tet1/2/3 heterozygote (Thet) mESCs were transplanted into foster mice and their ability to contribute to chimeras was determined. Embryos were dissected at E13.5 and revealed that approximately 60% of Thet cells contributed to chimeras, compared to 15% for TKO cells[254]. Interestingly, ectopic expression of Tet1 in TKO mESCs rescued the differentiation capacity and restored contribution to chimeras[254]. Consistent with this developmental impairment, mouse embryonic fibroblasts (MEFs) deleted in all 3 Tet paralogues could not be reprogrammed into iPSCs, inferring that oxidative DNA demethylation is required for pluripotent transition[255]. However, implementation of a CRISPR/Cas9 genetic deletion strategy to generate TKO mESCs, did not appear to affect pluripotency[256]. Consistently, TDG deficiency in mESCs also did not appear to disrupt pluripotency, suggesting active DNA demethylation is not required for ESC maintenance[204,203]. This is potentially supported by the observed embryonic lethality in

TDG-ablated mice at the E 11.5 developmental stage, which may therefore support a role of active demethylation in cellular differentiation[257].

Transcriptional analysis of TKO ESCs was associated gene activation and repression, attributable to hypermethylation at both enhancers and promoter regions[256]. Intriguingly, silenced genes upregulated in TKO cells were found to be primarily expressed in the early preimplantation embryo, indicative of stringent control of Tet activity in early embryogenesis[256]. Furthermore, inactivating all known Tet family members resulted in gastrulation defects, proposing the control of DNA methylation is fundamental for proper functioning of key signalling pathways to enable early body plan formation[258]. Further research is required to determine the exact functional role(s) of each Tet in development. However, owing to the murine neonatal lethality effects of a homozygous mutation in Tet3[242] it may be speculated, in agreement with the elevated mRNA expression profile during differentiation[252], that this paralogue is necessary for embryo development.

Staining of mouse embryos at the post-implantation stage signified the restriction of 5-hmC to multi-potent progenitor cells, and as development progressed further levels were found to decrease[249]. 5-hmC becomes restricted to terminally differentiated cell types, predominantly neuronal tissue and the bone marrow[249]. 5-hmC enrichment in the brain was established in 2009 through pioneering experiments in Purkinje neurons by TLC, high pressure liquid chromatography (HPLC) and mass-spectrometry techniques[259]. This has since been confirmed in multiple studies[182,260], and therefore suggests a role for this epigenetic mark in neuronal differentiation and neural plasticity[261]. Therefore, based on the literature evidence described thus far it can be alluded Tets/5-hmC are intrinsically linked to early developmental events. However, despite characterisation of the enzymatic function of Tet enzymes, investigations into the regulation of their activity, which may provide mechanistic insights underlying potential role(s) in cell specification events, remains largely unexplored.

1.6 *Chromatin dynamics are sensitive to [O_2]*

It has been recorded that 12% of the human population experience sustained ‘hypoxia’ as a consequence of low barometric pressure at high altitude[262]. This subsequently initiates as series of physiological adaptations that are regulated at the transcriptional level to maintain O_2 homeostasis, such as increased erythropoiesis, neovascularisation and metabolic reprogramming[263]. As stated previously, epigenetic modifications can regulate chromatin dynamics to enable TF accessibility to the DNA and thus may play a crucial role in the cellular response to changes in [O_2]. Hence, considering gradients of

O_2 form across the developing embryo, it is conceivable that O_2-dependent, 2-OGDO domain containing epigenetic modifiers influence chromatin structure and thus gene transcription.

Mechanisms that can facilitate alterations to the chromatin structure in response to changes to [O_2] include HMT and KDMs. Exposure of human embryonic kidney 293 cells (HEK cells) and human lung carcinoma A549 cells to 0.5% O_2 increased global levels of H3K9me2, which was attributed to elevated activity of the methyltransferase G9a[264]. G9a is essential for early embryogenesis *via* repression of developmental genes[265], and thus it can potentially be inferred that [O_2] in the embryo may define this HMT activity to ensure precise transcriptional regulation. Histone methylation marks can be removed by 2 different classes of KDMs, the lysine specific demethylase (LSD) and the JmjC domain containing demethylase[266]. The JmjC-KDMs, as alluded to previously, contain the 2-OGDO domain and are thus poised to respond to gradients of O_2. These demethylases are classified into 7 KDM subfamilies (KDM2-8) and can exhibit different histone substrate specificities, for examples KDM5 and KDM6 are specific for methylated H3K4 and H3K27 respectively[267]. Numerous studies on an array of cell types have directly demonstrated alterations to the histone methylation status in a variety of cells lines subjected to low O_2 tensions[267]. For example, a global increase in H3K4me2/3, H3K79me3, H3K9me2 and H3K27me3, was detected in Hepa1-6 cells after 48 hrs exposure to 0.2% O_2[268]. This result may be attributable to changes in JmjC-KDM activity and/or adaptions to their mRNA expression. For instance, a selection of JmjC-KDMs, such as KDM3A, KDM4C/D[269] and KDM6A[270], have been established as direct targets of HIF-signalling, and thus induction of their expression in response to low O_2 tensions may function as a compensatory effect for reduced KDM activity[271].

The absolute requirement of JmjC-KDMs for O_2 suggests these proteins may be O_2 sensors, and thus gives rise to the notion that [O_2] can be a determinant of chromatin structure[272]. Currently, biochemical data validating the sensitivities of JmjC-KDM to [O_2] are limited. Although K_m data is available for the KDM4 family, calculated for KDM4A, KDM4C and KDM4E as 57, 157 and 197 μM respectively[273,274] (further details of O_2 sensors are provided in Chapter 3: Results 1 Introduction). These differential O_2 sensitivities, which fall within the physiological [O_2] range, suggests O_2-signalling may differentially regulate KDMs in a spatial manner. This potentially may have consequences for gene regulation and therefore contribute to asymmetry in the developing embryo.

HIFs drive gene expression for an array targets when presented with a low [O_2] environment, which are in part dependent upon binding to its HAT domain containing CBP/p300 transcriptional coactivator[275]. O_2 has been shown to recruit p300 to erythropoietin (EPO) and VEGF genes to drive histone acetylation and thus gene transcription[276,277]. In fact, approximately 40% of HIF target genes are dependent upon this association with the HAT p300 or CBP[275,272]. Furthermore, HDAC activity and expression is also increased by low [O_2], which functionally has been shown to promote tumorigenesis by enhancing angiogenesis through suppression of hypoxia-responsive tumour suppressive genes[278]. This was further emphasised through inhibition of HDAC activity that blocked VEGF synthesis and resulted in decreased angiogenesis[279]. HDACs are also described to interact with HIF-1[70], for example HDAC7 was found to form a complex with HIF-1α to mediate nuclear translocation and thus transcriptional activity[280]. Overall, the influence of [O_2] on HATs and HDACs suggests that chromatin structure is tightly controlled, which is likely to function to enable precise transcriptional regulation. However, an additional layer of O_2-dependent transcriptional complexity was shown by the convergence of histone modifications to drive hypoxia-induced EMT. Here, HIF-1α-mediated activation of HDAC3 was shown to recruit H3K4-specific HMT activity to induce mesenchymal gene expression[281].

Similarly, Tet enzymes also contain the 2-OGDO domain, and thus it can be speculated that DNA methylation dynamics (and thus chromatin organisation) may also be regulated through [O_2]. However, at present investigation of Tet proteins in this context remain in its infancy and have not been considered in a developmental setting (more details are given in Chapter 3: Results 1). Although, independent of catalytic activity, Tet1 has been shown to function as a transcriptional co-activator, interacting with HIF-2α to regulate hypoxia-responsive gene expression and EMT[282]. By contrast, from cancer studies, DNA methylation has shown to be sensitive to changes in [O_2][283]. Exposure of human colorectal and melanoma cell lines to low [O_2] displayed a reduction in 5-mC[284]. Similarly, in a human hepatoma cell line, a low O_2 tension decreased the level of the methyl donor SAM, which culminated in DNA hypomethylation[285]. This tumour associated reduction in 5-mC has been shown to facilitate the binding of HIF-1 to the HRE and thus augment the HIF-mediated effects on malignant cell growth[286]. Although, prolonged exposure (14 weeks) to low O_2 tensions was shown to increase 5-mC levels in a PwR-1E cells, a prostate epithelial line[287]. It was further demonstrated that this hypermethylation was not a result of increased Dnmt activity, but instead may be due to a specific increase in Dnmt3b mRNA expression[287]. In addition, the mRNA expression of Dnmt enzymes were confirmed to be regulated by [O_2], but were reported to be increased in cardiac fibroblasts and downregulated in colorectal cancer cells[288,289]. Collectively, owing to the sensitivity

to 5-mC by O_2, it is reasonable to believe that demethylation may be integrated into regulating these methylation changes.

1.7 *Hypothesis*

Tet enzymes are O_2 sensors at concentrations deemed physiologically relevant in embryogenesis. The catalytic activity of Tet1, Tet2 and Tet3 is differentially regulated by [O_2] to influence gene transcription and thus cell fate decisions, thereby determining embryonic asymmetry.

This hypothesis will be tested over 3 independent results chapters, which each contain their own detailed aims.

Chapter 2: Materials and Methods

2.1 *Reagents*

All reagents used in these studies were purchased from Sigma unless otherwise stated. Primers were purchased from Sigma or Integrated DNA Technologies.

2.2 *Cell Culture*

All cell cultures, unless stated, were conducted in a humidified 37 °C incubator exposed to a gas mixture of 5% CO_2 and 95% air.

2.2.1 Murine P19 embryonal carcinoma cells

Mouse P19 embryonal carcinoma cells (a kind gift from Mona Nema) were cultured in α-Minimum Essential Medium Eagle (Sigma, M8042) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and 1% L-glutamine-penicillin-streptomycin solution. Cells were cryopreserved in FBS containing 10% DMSO in a Coolcell® freezing container (Biocision) at -80 °C overnight, this ensured a cool rate of -1 °C/minute. Cryovials were transferred to liquid nitrogen for long-term storage.

P19 cells were induced to differentiate into EB-like aggregates through culture onto 100 mm^2 non-adherent petri dishes at a density of $1x10^6$ cells/dish. Differentiation was directed towards a cardiac and neural-like fate in the presence of 1 μM retinoic acid (RA) and 1% dimethyl sulfoxide (DMSO) respectively. Medium was changed every 3 days by pelleting EB-like aggregates under gentle centrifugation (800 G, 3 mins).

2.2.2 Mouse embryonic stem cells

R1 mESCs (ATCC®), provided by Shukry Habib (Centre for Stem Cell & Regenerative Medicine, King's College London), were maintained undifferentiated on 0.1% gelatin coated flasks in Advanced DMEM/F12 (Invitrogen, 12634) containing 8 μM 2-mercaptomethanol (Invitrogen), 2 mM GlutaMAX-I™ (Invitrogen), 100 U/ml penicillin and 100 μg/ml streptomycin (Invitrogen). Medium was further supplemented with 10% EmbryoMax® FBS (Millipore), 20 ng/ml leukemia inhibitory factor (LIF) (Millipore), 3 μM CHIR-99021, 1 μM PD0325901 and 5 μg/ml plasmocin prophylactic (InvivoGen). Cell medium was changed daily and cells passaged with StemPro Accutase (Invitrogen) every 2 days in a 1:4 ratio. Cells were cryopreserved in EmbryoMax® FBS containing 10% DMSO, as described above.

mESCs were induced to differentiate unbiasedly into EBs by plating cells onto 100 mm^2 non-adherent petri dishes, at a density of $1x10^6$ cells/dish. Cells were differentiated in KnockOut DMEM (Invitrogen, 10829) containing 15% KnockOut™ (KO) serum replacement (Invitrogen), 0.1 mM MEM amino acid solution (Invitrogen) and 2 mM GlutaMAX-I™. Media was changed every other day by pelleting EBs by gentle centrifugation (800 G, 3 mins).

2.2.3 Human embryonic kidney cells, rat cardiomyoblasts and human neuroblastoma cells

HEK-293T (ATCC®), rat cardiomyoblasts (H9c2) (ATCC®) and human neuroblastoma (SH-SY5Y) (provided by John Pizzey, Guy's Campus, King's College London) cells were cultured in Dulbecco's Modified Eagle's medium (high glucose) (Sigma, D6546) complemented with 10% heat-inactivated FBS and 1% L-glutamine-penicillin-streptomycin solution. Cells were cryopreserved in FBS containing 10% DMSO as described above.

2.3 *Low O_2 culture*

Cells were exposed to low O_2 tensions using a ProOx C21 regulated C-chamber (BioSpherix) positioned inside a humidified 37 °C incubator. Desired O_2 tensions were achieved by the control of N_2 balance in 5% CO_2. For low O_2 cellular differentiation experiments, fresh medium was equilibrated to the desired O_2 tension overnight prior to medium changes. To determine the ability of the O_2 chamber to activate hypoxic signalling pathways, mESCs were plated onto 6-well plates at a density of 250,000 cells/well and 24 hrs later exposed to 3% O_2 for 4 and 24 hrs. Protein was harvested and probed for HIF-1α levels, as described in section 2.10.

2.4 *Tet overexpression*

Bacteria containing Tet overexpression plasmids, were kindly provided by Anjana Rao (Addgene plasmids: Tet1; 49792[168,290,291,292,293], Tet2; 41710, Tet3; 49446). These were streaked onto agar plates from the bacterial stabs provided. Single colonies were isolated, inoculated into lysogeny broth (LB), and cultures grown overnight at 37 °C under agitation. Plasmid DNA was purified using a HiSpeed plasmid maxi kit (Qiagen) and stored at -20 °C.

To optimise transfection conditions for Tet overexpression in mESCs, an array of transfection reagents were tested for their transfection efficiency using a GFP expression plasmid (Invitrogen). According to their manufacturer's instructions: Viafect™ (Promega), Lipofectamine® 2000 (Invitrogen), TurboFect™ (Thermo Scientifc), FuGENE® HD (Promega) and Xfect™(Clontech) were used to transfect 1 μg of GFP plasmid into 24-well plates pre-seeded with 100,000 mESCs/well 24 hrs before. In addition, nucleofection was also used as alternative transfection technique to further assess the transfectability of mESCs. A 4D-nucleofector™ system (Lonza), using the P3 primary cell kit was used to optimise the transfection of a pmaxGFP™ plasmid (Lonza), in accordance with manufacturer's instructions. 2 μg pmaxGFP™ plasmid was combined with $3x10^6$ mESCs suspended in Nucleofector™ solution containing Nucleofector™ supplement, in a total volume of 100 μl. The reaction mix was transferred into Nucleocuvette™ vessels and exposed to either one of 4 pulse programs: CG-104, CA-137, CB-150, CD-118 within the 4D-nucleofector™ system. Following re-suspension in 500 μl of fresh medium, cells were plated into 24-well plates at 100,000 cells/well. GFP expression was analysed 24 hr post-transfection on a Leica fluorescent microscope and cell images acquired with a Q-imaging® camera and Openlab™ software (PerkinElmer).

Despite success with GFP transfection, Tet proteins were not expressed under the same transfection conditions. Therefore, HEK cells were used as an overexpression model for future studies. HEK cells were plated onto 6-well plates at 400,000 cells/well in 3 ml of medium, 24-hrs prior to transfection. Transfection mixes for overexpression of each Tet, and pcDNA™ 3.1 (Invitrogen) transfection control, was conducted in triplicate as follows: 6 μg of plasmid DNA was added to 300 μl Opti-MEM™ (Invitrogen) and mixed with 300 μl Opti-MEM™ containing 12 μl of Lipofectamine® 2000. Following a 10 min incubation at room temperature, 200 μl of transfection mix was added to each well and cells harvested for downstream applications after 48 hrs. For low O_2 experiments, HEK cells were exposed to the desired O_2 tension directly after transfection mixes were added to the cells.

Tet proteins were also overexpressed in P19 cells for a microarray study (see section 2.15). P19 cells were plated onto T25 culture flasks at 500,000 cells/flask in 5 ml of medium, 24 hrs prior to transfection. For a single transfection, 15 μg of either Tet overexpression construct (or pcDNA™ 3.1 control) was mixed with 45 μl of Fugene® HD transfection reagent in a total of 500 μl of Opti-MEM™. Following a 15 min incubation at room temperature, DNA and transfection reagent complexes were added to respective culture flasks.

2.5 *Cell treatments*

2.5.1 Dimethyloxalyglycine and colbalt chloride

The ability to activate HIF-dependent signalling in P19 cells chemically by Dimethyloxalylglycine (DMOG) was assessed. P19 cells were plated onto 6-well culture plates at 250,000 cells/well in 3 ml of medium and 24 hrs later treated with 0.5 mM and 1 mM DMOG for 16 hrs. Cells were harvested and protein levels of HIF-1α determined, as described in section 2.10. mESCs were treated with 100 μM cobalt chloride (a less toxic alternative to DMOG) for 24 hrs as a positive control for the induction of HIF-1α protein for assessing the function of the O_2-controlled chamber, see section 2.3.

To understand whether chemical activation of O_2-dependent signalling could influence cellular differentiation, P19 cells were differentiated into EB-like aggregates in the presence of 1 mM DMOG for 4 days and expression of lineage markers determined by QPCR analysis, see section 2.7. In a separate preconditioning experiment, P19 cells were differentiated in the presence of 1 mM DMOG for 4 days, before being allowed to differentiate further till days 7 and 11. The effect of this initial hypoxic-mimetic stimulus upon lineage marker expression at later differentiation time points was assessed by QPCR.

2.5.2 Ascorbic Acid and inhibitors

mESCs were plated onto 6-well culture plates at a density of 250,000 cells/well in 3 ml of medium, 24 hrs prior to the addition of AA (vitamin C) at increasing doses (0.01 mM, 0.1 mM, 1 mM, 5 mM and 10 mM) (AA solution was made up in water and neutralised to pH 7.2 using 5 M sodium hydroxide). Cells were harvested 24 hrs after treatment and 5-hmC levels detected by a dot-blot, as described in 2.8. Differentiating mESCs were also treated with 1 mM AA and differentiated for 3 days. Cells were subsequently harvested and 5-hmC levels detected as described in 2.8.

To investigate whether 5-hmC levels were affected by a reduction in intracellular AA, the use of phloretin and buthionine sulfoximine (BSO), inhibitors of AA uptake transporters and dehydroascorbic acid (DHA) reduction, respectively, were studied. mESCs were plated onto 6-well plates at 250,000 cells/well in 3 ml of medium and 24 hrs later treated with 1 mM BSO (made up in water) or 200 μM phloretin (made up in DMSO). In addition, cells were pre-treated with 1 mM BSO or 200 μM phloretin for 3 hrs, before addition of 1

mM AA. Cells were harvested for 5-hmC analysis after 24 hrs as described in section 2.8. Furthermore, the effect of these inhibitors upon 5-hmC levels during mESC differentiation was also examined. mESCs were differentiated in the presence of increasing doses of BSO (100 μM, 500 μM and 1 mM) and phloretin (100 μM and 200 μM) for 3 days before cells were harvested for 5-hmC analysis, as described in section 2.8. For phloretin treatments, a DMSO vehicle control was included at an equivalent DMSO level to that contained in the 200 μM dose.

Owing to the different AA concentrations within mESC maintenance and differentiation medium, both cell mediums were tested to ensure the 5-hmC profile of differentiating mESCs was not a result of changes in AA levels. Firstly, undifferentiated mESCs were plated at 250,000 cells/well into 6-well culture dishes and cultured in either fresh maintenance medium or maintenance medium containing 15% KO serum (used in differentiation studies) in replace of EmbryoMax® FBS. As a control, 1 mM AA treatment was included in this study. Cells were harvested 24 hrs later for 5-hmC analysis, see section 2.8. In an independent investigation, mESCs were plated into T25 culture flasks at 625,000 cells/well and pre-treated with different serum-containing maintenance medium for 24 hrs, as described above, before being induced to differentiate in full differentiating medium (containing 15% KO serum) for 3 days (described in 2.2.2). 5-hmC levels were then quantified as described in section 2.8.

2.6 *Generation of stable Tet ablated mESC lines*

Genetic ablation of each Tet in mESCs was conducted using MISSION® short hairpin RNA (shRNA) lentiviral transduction particles. Lentiviruses are capable of infecting dividing and non-dividing cells and stably integrating into the host genome resulting in long term transgene expression. The shRNA is transcribed by the pol III promoter U6 from the pLKO.1-puro vector into 2 complementary 19-22 base pair (bp) RNA sequences, linked by a short loop of 4-11 nucleotides[294]. This shRNA molecule is exported to the cytosol where it is recognised by Dicer and processed into siRNA duplexes. These siRNA duplexes are incorporated into the RNA-induced silencing complex (RISC) and subsequently targeted to specific mRNA molecules initiating their cleavage and degradation by Argonaute 2 (Ago2)[294].

2.6.1 Puromycin dose response curve

To select for positively transduced cells, the sensitivity of mESCs to puromycin (a resistance gene is contained within the lentiviral vectors) was determined by a dose response curve. mESCs were seeded onto 0.1% gelatin coated 24-well tissue culture plates at 50,000 cells/well, 24 hrs prior to the addition of increasing concentration of puromycin (0.5-10 μg/ml) (Invitrogen). Puromycin containing media was changed daily. 1 μg/ml, the lowest concentration to induce cell death after 72 hrs, was used to select for positive lentiviral transduced mESCs.

2.6.2 Lentiviral transduction

mESCs were seeded onto 0.1% gelatin-coated 24-well tissue culture plates at 20,000 cells/well in 1 ml of medium, 24 hrs prior to lentiviral transduction. Lentiviral transduction particles were thawed on ice and the volume of particles to give a multiplicity of infection (MOI) of 15 (see equations in Figure 2:1) were added to culture medium, containing 10 μg/ml hexadimethrine bromide, to give a total volume of 250 μl. For each transduction, mESC medium was replaced with 250 μl lentiviral containing medium. 2 clones for each Tet paralogue were available: Tet1 (TRCN0000341849 and TRCN0000341850), Tet2 (TRCN0000192770 and TRCN0000201087) and Tet3 (TRCN0000376843 and TRCN0000375340). A non-targeted shRNA (MISSION®, pLKO.1, (SHC016V-1)) control cell line was also generated to control for Tet silencing studies. Virus containing medium was incubated on cells overnight and subsequently replaced with 500 μl fresh medium. 72 hrs post-infection transduced mESCs were cultured in medium containing 1 μg/ml puromycin. Selected cells were expanded and subsequently cryopreserved as described in 2.2.2. Successful ablation of the Tet enzymes, compared to shRNA control, was assessed at the mRNA and protein level by QPCR and Western blotting respectively.

MOI x cell number = TU

TU/TU per ml = ml of lentiviral particles required

Figure 2:1 Equations required for calculating the volume of lenti-viral shRNA particles for a desired multiplicity of infection for mESC transduction

2.7 *Quantification of mRNA expression*

2.7.1 RNA isolation

RNA was extracted from cell and tissue using ReliaPrep™ RNA tissue miniprep system (Promega), according to manufacturer's instructions, but modified by the addition of 1 μl RQ1 DNase (Promega) per sample during the DNase incubation step. RNA was quantified and assessed for purity using a NanoDrop spectrophotometer (ND-1000, Labtech international). RNA was stored at -80 °C. Note, for overexpression studies an additional in-solution DNase step, to ensure all plasmid DNA was removed, was performed on purified RNA, as described for the gene profiler arrays in section 2.7.2.

Differentiating P19 and mESCs were harvested by pelleting cell aggregates under gentle centrifugation (800 G, 3 mins), washed in ice cold phosphate buffered saline (PBS), pelleted and lysed in 500 μl RNA lysis buffer (LBA) containing 1-Thioglyerol (Promega). Adherent cells were harvested from tissue-culture plates, after washing with ice cold PBS, by scraping into LBA (volume varied depending on the size of the culture dish, but did not exceed 500 μl). If required, cell lysates were centrifuged through Qiashredder columns (Qiagen) (13000 G, 1 min).

Tissue samples (20 mg) were lysed in 500 μl of LBA and homogenised using FastPrep® lysing beads D (MP biomedicals) on a Precellys 24 machine (Bertin) at program 3 for 20 sec bursts (usually performed twice).

2.7.2 cDNA synthesis

cDNA synthesis from a single stranded RNA template occurs *via* a process called reverse transcription. Here the polyadenylated 3' end of the mRNA becomes hybridised to a oligodeoxythymine primer (Oligo dT_{18}), and extended with a reverse transcriptase enzyme in the presence of dNTPS to yield a single stranded cDNA copy.

For each RNA sample, 1 μg was diluted in RNase/DNase-free water to a total volume of 13 μl, and mixed with 1 μl oligo-dt_{18} primers (1 μg/ml) and 1 μl dNTPs (10 mM) (Promega). Samples were initially heated at 70 °C for 3 mins to disrupt any secondary structures that may be present within the RNA and incubated at 4 °C. 4 μl of 5X reverse transcriptase buffer (Promega), 0.5 μl RNase inhibitor (Promega) and 0.5 μl reverse transcriptase (Moloney murine leukaemia virus) (Promega) were then added, giving a

total reaction volume of 20 μl. Negative controls comprising all reaction components, except the reverse transcriptase enzyme (substituted with 0.5 μl RNase/DNase-free water) were used to assess for genomic DNA contamination within the sample. All samples were incubated at 42 °C for 90 mins and heated to 70 °C for 10 mins. cDNA was diluted in 80 μl of RNase/DNase-free water and stored at -20 °C.

For mouse cell lineage gene profiler PCR arrays (RT^2 profiler PCR array, 384-well format (Qiagen)), RNA samples were treated with an additional in-solution DNase step as recommended by the manufacturer's instructions prior to reverse transcription. Each DNase reaction contained the RNA sample (1-2 μg), 1 μl RQ1 10X reaction buffer (Promega) and 1.5 μl RQ1 DNase (Promega) made up to a total volume of 10 μl in RNase/DNase-free water. Samples were heated at 37 °C for 30 mins. The reaction was terminated by the addition of 1 μl RQ1 DNase stop solution (Promega) and a 65 °C heat step (for 10 mins), to inactivate the DNase. Triplicate RNA samples were equally pooled, and a total of 400 ng of RNA was reversed transcribed, as conducted above, and diluted with 91 μl RNase/DNase-free water (giving a total volume of 111 μl).

2.7.3 Quantitative PCR (QPCR)

Quantitative PCR (QPCR) was used to determine the relative mRNA expression of specific genes. This analysis method relies on fluorescence detection. For experiments described here, SYBR green, a nucleic acid stain, emits a fluorescent signal when bound to double stranded DNA by intercalating with DNA bases. Therefore, the fluorescent signal increases proportionally to the amount of double-stranded DNA product amplified after each PCR cycle, enabling quantitation of the gene of interest's expression across sample sets.

Each QPCR reaction was performed in a 20 μl reaction volume, made up as follows: 2 μl cDNA, 2 μl of 3 μM gene-specific primers (see primer list in Table 3) 6 μl RNase/DNase-free water and 10 μl 2X qPCRbio Sygreen mix Hi-ROX (PCR Biosystems). QPCR reactions were conducted on a StepOnePlus™ QPCR machine (Applied Biosystems). Samples were initially denatured (95 °C, 10 mins) and amplified for 40 cycles at the following conditions: denatured at 95 °C, for 15 secs, annealed at 60 °C for 30 secs and extended at 72 °C for 20 secs. Melt curve analysis was performed for all reactions in order to determine primer specificity, see section 2.7.4.

For the gene profiler PCR arrays, QPCR reagents were prepared as follows: 102 μl of cDNA, 650 μl 2X qPCRbio Sygreen mix Hi-ROX and 548 μl of RNase/DNase-free water. 10 μl of QPCR mix was dispensed by a Bravo automated liquid handling platform (Agilent) to each well. QPCR reactions were conducted on a Viia™ 7 system (Applied Biosystems), under the following conditions: initial denaturation step (95 °C, 10 mins) and 40 cycles of denaturation (95 °C, for 15 secs) and annealing (60 °C for 1 min). Melt curve analysis was performed for all reactions in order to determine primer specificity.

Quantification of mRNA expression was determined from the amplification plot of natural log fluorescence against cycle number. A cycle threshold (C_t) was determined from the linear phase of the exponential amplification curve for each tested gene. Hence, the C_t of each sample was determined as the cycle number at which fluorescence reached this threshold. Relative mRNA expression was quantified using the comparative C_t ($\Delta\Delta C_t$) method. The target gene of interest was normalised to a control gene (see 2.7.4) and subsequently normalised to a control sample, as shown from the equation below (Figure 2:2). The expression of the target gene for each sample relative to control is then calculated as $2^{-\Delta\Delta Ct}$.

$$\Delta\Delta C_t = (C_{t\ \text{target (sample)}} - C_{t\ \text{normalisation control (sample)}}) - (C_{t\ \text{target (control sample)}} - C_{t\ \text{normalisation control (control sample)}})$$

Figure 2:2 Comparative C_t method for relative quantification of mRNA expression. Relative mRNA expression for each sample against the tested target gene was calculated using the $\Delta\Delta C_t$ method as shown.

2.7.4 QPCR primer design

Gene specific primers were designed to bridge exon-exon junctions using the NCBI primer design tool (www.ncbi.nlm.nih.gov/tools/primer-blast/). Primers were designed to have a minimal and maximal amplicon size of 70 and 200 bp respectively, with an optimal melting temperature of 60 °C. The specificity of primers was assessed for their ability to produce a solo melting peak and further confirmed through the detection of a single PCR product on a DNA gel by electrophoresis (visualised by Nancy-520 nucleic acid stain), an example is shown in Figure 2:3. A table of validated QPCR primer sets are displayed in Table 3.

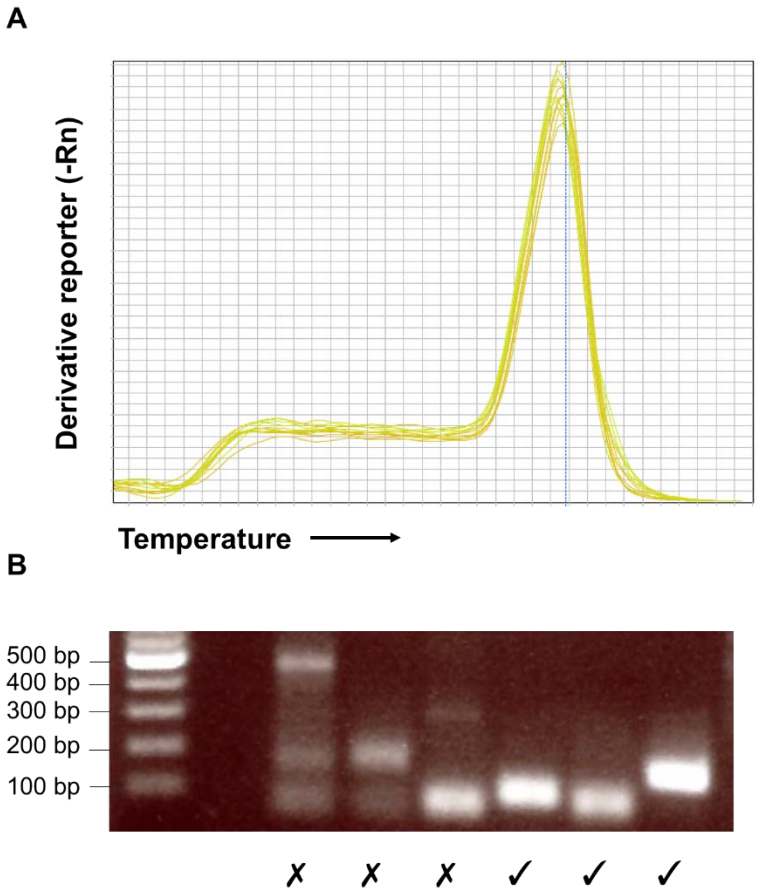


Figure 2:3 QPCR Primer validation. (A) Melt curve analysis performed after each QPCR run to confirm amplification of a single product (depicted by a single melt peak). (B) Primer sets were further verified to produce a single gene product by gel electrophoresis after QPCR amplification. Primer pairs producing multiple bands were discarded from future studies, denoted by the cross, and redesigned.

A normalisation control gene was selected from the geNorm™ reference selection kit (Primerdesign) in accordance with manufacturer's instructions. The mRNA expression of an array of normalisation genes were assessed against mESCs samples at their undifferentiated and day 3-, day 7- and day 11-differentiated state. Expression data was analysed using geNorm qbase software to determine the most stably expressed genes among the tested sample sets. Canx was selected as a normalisation control gene for QPCR experiments.

2.8 *Dot-blotting*

The dot-blot technique was used to immobilise DNA onto nylon membranes in a series of spots, which is amenable for densitometry analysis to assess global changes in 5-mC and 5-hmC levels.

2.8.1 DNA isolation

EBs were harvested by gentle centrifugation (800 G, 3 mins), washed in ice cold PBS, pelleted and stored at -20 °C. Adherent cells were washed and scraped into ice cold PBS, transferred to a microcentrifuge tube, centrifuged (800 G, 5 min) and cell pellets stored at -20 °C.

Genomic DNA was prepared from cell pellets by 350 µg/ml proteinase K (Sigma) digestion in 15 mM NaCl, 1% Sodium dodecyl sulfate (SDS), 100 mM Ethylenediaminetetraacetic acid (EDTA), 50 mM Tris-HCl pH 8 in a total volume of 700 µl, at 55 °C for 4 hrs for cells and overnight for tissue (samples up to 40 mg). Samples were sonicated (Branson 150 sonfier, setting 2) twice for 10 secs to obtain 200-1000 bp fragments, before undergoing RNaseA (140 µg/ml) (Qiagen) treatment for 30 mins at 55 °C. DNA was phenol/chloroform extracted using 5 prime phase lock gel tubes (Scientific Laboratory Supplies) and ethanol precipitated. DNA pellets were re-suspended in up to 200 µl of RNase/DNase-free water and quantified using a NanoDrop spectrophotometer.

2.8.2 Immunoblotting

2 µg of DNA was applied to a Hybond™-N membrane (GE Healthcare) using a dot-blot hybridisation manifold (Cleaver Scientific), as described by Brown, T (2001)[295]. DNA was prepared in a total volume of 105 µl containing 45 µl 20X SSC (3 M NaCl and 300 mM trisodium citrate, adjusted to pH 7) and made up to volume in RNase/DNase-free water. Known DNA standards for 5-mC (10 ng) (Actif Motif) and 5-hmC (5 ng) (Actif Motif) were included in initial optimisation experiments as separate samples to control for antibody specificity. DNA was denatured at 100 °C for 10 mins and placed on ice. In parallel, membranes were soaked 6X SSC and incubated for 10 mins. The dot-blot manifold was assembled with an initial filter paper layer, pre-soaked in 6X SSC, with the membrane placed on top, secured firmly and 0.8 bar pressure applied. An initial 150 µl of 6X SSC was applied to each well prior to addition of the 150 µl of the DNA sample. Once samples had passed through, the manifold was dismantled and the membrane submerged in denaturation solution (1.5 M NaCl, 0.5 M NaOH) for 10 mins followed by neutralisation

solution (1M NaCl, 0.5 M Tris-HCl) for 5 mins. Membranes were dried and exposed to UV irradiation on a UVP GelDoc-it imager (5 mins) to covalently bind DNA to the nylon membrane. Next, membranes were blocked in 10% milk (Marvel) in PBS with 0.1% Tween-20 (PBS/T) for 5-hmC blots or with Odyssey® blocking buffer (PBS) (LiCor) for 5-mC blots, for a minimum of 1 hr at room temperature. Membranes were probed with 5-hmC and 5-mC antibodies diluted in their respective blocking conditions (see Table 2 for antibody information) overnight, under agitation, at 4 °C. Subsequently, blots were washed 3 times for a minimum of 10 mins in PBS/T, and incubated with IRDye® 800CW secondary antibodies (LiCor) (diluted 1:15000) (5-mC (LiCor, 926-32210) and 5-hmC (LiCor 926-32213)) in respective blocking conditions, shielded from light, for 1 hr under agitation at room temperature. Membranes were then washed 3 times for a minimum of 10 mins in PBS/T, with an additional final PBS-only wash for 5 mins prior to signal detection on an Odyssey® Clx imaging system (LiCor). Blots were quantified using Image Studio™ software (LiCor), enabling the ratio of 5-hmC to 5-mC to be calculated for each sample.

2.9 *Mass spectrometry analysis*

Mass spectrometry was utilised as a sensitive and accurate method to quantify and distinguish between 3 nucleotide bases of interest; C, 5-mC and 5-hmC, within the same sample.

Genomic DNA was prepared for mass-spectrometry analysis in accordance with Le, T *et al* (2011)[296]. DNA was broken down to individual nucleoside components as follows: 1 µg DNA was mixed with 2.5 µl 10X DNA Degradase™ reaction buffer (Zymo Research) and 1 µl DNA Degradase Plus™ (Zymo Research) made up to a total volume of 25 µl with RNase/DNase-free water. Samples were incubated at 37 °C for 4 hrs and the reaction inactivated by the addition 175 µl of 0.1% formic acid to yield a final concentration of 5 ng/µl of DNA.

DNA samples were diluted 4 times in 25 µl DNA Degradase™ reaction buffer and 175 µl of 0.1% formic acid before injection into an LC column (Zorbax Eclipse plus C18 RR HT 100 x 2.1mm, 1.8µ particle size) integrated to an Agilent 1100 LC system, which was set to a flow rate of 0.1 ml/min of water/methanol/formic acid (95%/ 5%/ 0.1%). The effluent of the column was directed to electrospray ion source interfaced with Waters Quattro LC triple quadrupole mass spectrometer. The intensity of specific MH^+ > fragment ion transitions were recorded (C m/z 228.1→112.1, 5-mC m/z 242.1→126.1 and 5-hmC 258.1→142.1). The multiple reaction monitoring (MRM) peak area was

identified for C, 5-mC and 5-hmC, shown in Figure 2:4A/B. A series of DNA standards of C, 5-mC and 5-hmC (Actif Motif) were prepared for calibration at a concentration of 1 µg/ml in MeOH, mixed and diluted to 1.5 ng/µl in 25 µl DNA Degradase™ reaction buffer and 175 µl of 0.1% formic acid (Figure 2:4A). Amounts of 5-mC and 5-hmC were normalised to total C (C, 5-mC, and 5-hmC) within each sample.

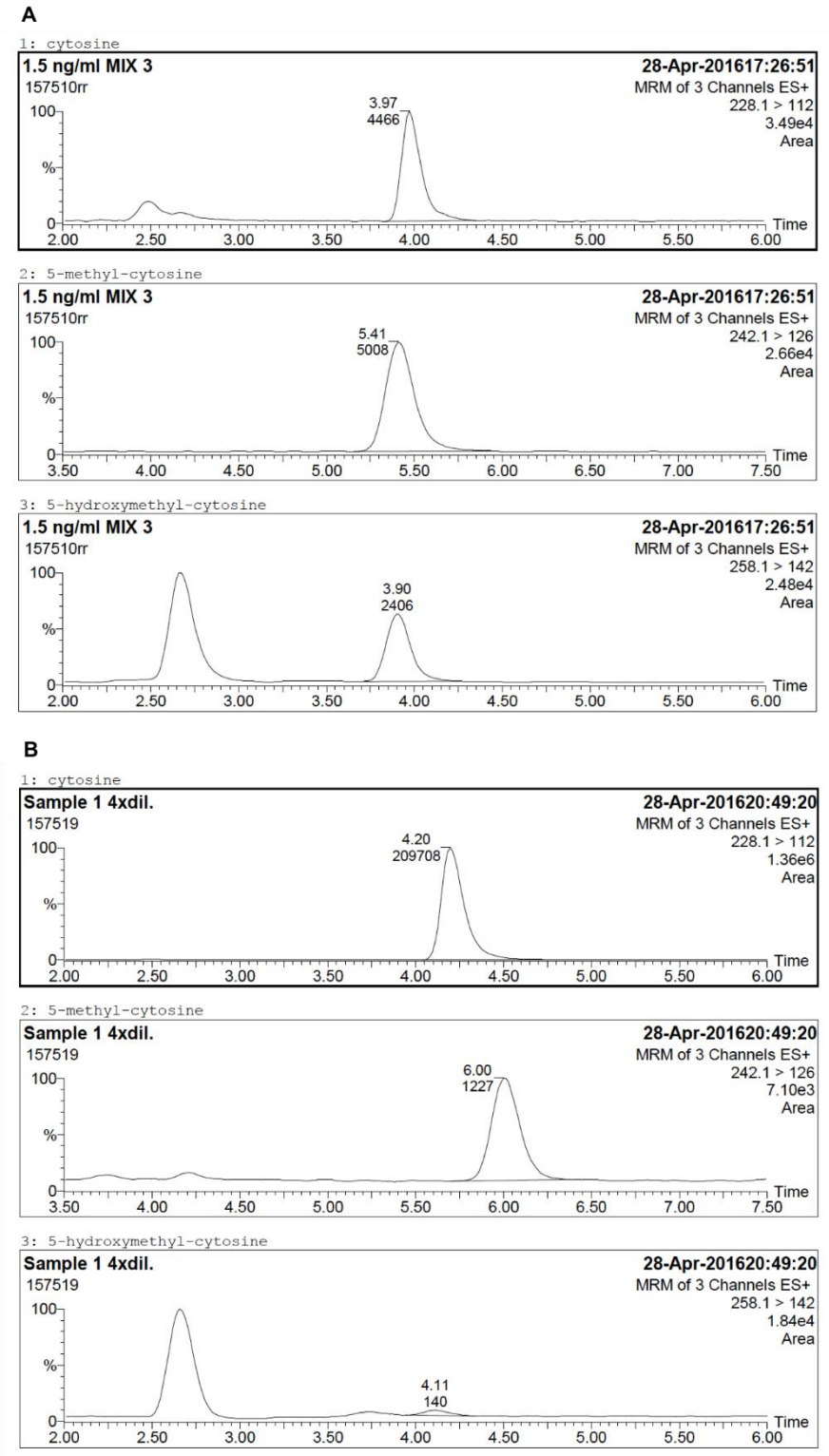


Figure 2:4 Mass-spectrometry chromatograms for C, 5-mC and 5-hmC detection. (A) Chromatograms showing the detection of 1.5 ng/μl of C, 5-mC and 5-hmC DNA standards. (B) An example of a set of chromatograms for C, 5-mC and 5-hmC detection in an undifferentiated mESC sample.

2.10 *Western blotting*

Western blotting is a technique used to analyse the expression of specific proteins within a sample by polyacrylamide gel electrophoresis, which relies on the separation of proteins according to their molecular weight.

2.10.1 Sample preparation

Cells were washed and scraped into ice cold PBS, transferred into a microcentrifuge tube and pelleted by gentle centrifugation (800 G, 3 mins). Cell pellets were either stored at -80 °C or lysed, dependent on pellet size in up to 300 µl of protein lysis buffer (150 mM NaCl, 2 mM EGTA, 5 mM EDTA, 30 mM sodium fluoride, 40 mM β-glycerophosphate, 20 mM sodium pyrophosphate, 1 mM sodium orthovanadate, 1 mM phyenylmethylsulfonyl-fluroide, 0.5% Nonidet P-40, in 25 mM Tris-HCl at pH 7.4, supplemented with 5 µl/ml protease inhibitor cocktail). In some instance, where lysates appeared insoluble, samples were briefly sonicated for 10 secs (Branson 150 sonfier, setting 1).

Protein concentration was determined by a Bradford colorimetric protein assay, which utilises the spectral properties of Coomassie Brilliant Blue G-250 upon protein binding. A standard curve, performed in triplicate, was made with increasing concentrations of bovine serum albumen (BSA) (0-50 µg/ml) supplemented with 2 µl of protein lysis buffer (since Bradford reagent is sensitive to detergents in lysis buffer) made up to a total volume of 200 µl with Bradford reagent. 2 µl of each protein lysate was added to 198 µl of Bradford reagent in triplicate and after a short incubation absorbance measured at a wavelength of 595 nm (Tecan, GENios pro). Sample protein content was determined from the standard curve and added to 5X loading buffer (0.5 M DTT, 0.02% Bromophenol blue, 30% glycerol, 10% SDS in 250 mM Tris-base, pH 6.8), heated at 95 °C for 10 mins (to denature proteins) and stored at -20 °C.

2.10.2 Immunoblotting

Polyacrylamide gels were cast (see Table 1 for volume and reagents) and 30 µg of protein sample was loaded to each well alongside a molecular weight marker (Bio-rad). Gels were electrophoresed in running buffer (25 mM Tris-base, 192 mM Glycine, 0.1% SDS) at 100 V through the stacking gel, before increasing to 150 V through the resolving gel. Proteins were transferred onto nitrocellulose membranes at 100 V for 70 mins in transfer buffer (25 mM Tris-base, 192 mM Glycine, 10% methanol). Membranes were

either stained with Ponceau S red to confirm efficient transfer, or alternatively total protein/lane was quantified using Revert total protein stain kit (LiCor), according to manufacturer's instructions.

Membranes were blocked in either 10% milk Tris-buffered salt solution containing 0.1% tween-20 (TBS/T) or 10% milk PBS/T for a minimum of 1 hr at room temperature. Blots were probed with primary antibodies overnight at 4 °C under agitation (see Table 2 for primary antibodies used, complete with conditions). Membranes were washed in TBS/T or PBS/T 3 times for a minimum of 10 mins and incubated with the appropriate IRDye® 800CW (926-32213/926-32212) or IRDye® 680RD (926-68072) secondary antibodies (1:15000) for 1 hr at room temperature, under agitation. Following 3 washes for a minimum of 10 mins in TBS/T/PBS/T and a final wash in either TBS/PBS alone, blots were visualised on an Odyssey® Clx imaging system.

Blots were quantified using Image Studio™ software. Signal from the protein of interest was normalised to the signal of a loading control gene or the total protein signal/lane.

2.11 *^{1}H-NMR metabolite assessment*

HEK cells were plated onto 150 mm^2 tissue culture dishes at a density of $4x10^6$ cells/dish for atmospheric O_2 and 6.5 $x10^6$ cells/dish for low O_2 studies. After 24 hrs, cells were exposed to 3% and 1% O_2 (or kept under atmospheric conditions) for 24 hrs. Cells were washed in ice cold PBS twice and incubated on dry ice to quench metabolites for 3 mins. Next, cells were scraped into 1.7 ml of methanol (a 10 µl aliquot was taken for protein quantification), mixed with 1.7 ml chloroform and incubated under agitation at 4 °C for 10 mins. 1.7 ml of water (Milli-Q) was added to the metabolite extraction buffer and centrifuged (800 G, 45 mins, 4 °C). The top (polar) fraction was collected and the methanol and water evaporated using a SpeedVac™ concentrator (Thermo Scientific) at 30 °C. Dried extracts were reconstituted in a 100 mM sodium monophosphate buffer (pH 7.0) containing 500 µM sodium 3-(trimethylsilyl)propionate-2,2,3,3-d4 (TMSP), 1.5 mM sodium azide , 0.5 mM EDTA and 100% deuterium oxide (D_2O).

A 700 MHz Bruker spectrometer equipped with a cryogenic probe was used for 1D ^{1}H NMR data acquisition with supressed water resonance. ^{1}H NMR spectra were acquired using 9.3 kHz spectral width and 32 K data points with acquisition time of 1.67 secs, relaxation delay of 5 secs and 128 scans. The resulting spectra were processed to 65536 data point and corrected for phasing and zero referencing using NMRLab[297]. Resonance assignments and quantification were made with reference to Chenomx NMR Suite 7.1

(Chenomx). Data was normalised to total protein, determined by a Bradford assay, as described 2.10.1.

2.12 *5-hmC-DNA immunoprecipitation-sequencing/PCR*

The 5-hmC-DNA immunoprecipitation (hMeDIP) (Active Motif) assay was used to enrich for the 5-hmC mark on genomic DNA using a 5-hmC antibody. These enriched DNA fractions were analysed by a non-targeted genome wide sequencing approach (hMeDIP-seq) and a more targeted PCR approach (hMeDIP-PCR).

Genomic DNA was isolated from day 3-differentiated mESCs, exposed to atmospheric and 1% O_2, as described in section 2.8.1. Triplicate samples were pooled equally and a total of 1 µg of fragmented DNA (Figure 2:5) was used for the hMeDIP protocol in accordance with manufacturer's instructions. In brief, fragmented DNA was incubated in a reaction mix containing 4 µl 5-hmC antibody (or IgG control) in a total volume of 100 µl, overnight at 4 °C with rotation. An aliquot of fragmented DNA was set aside as an input control for downstream applications. Protein G magnetic beads were added to each immunoprecipitation (IP) and a further incubation of 2 hrs at 4 °C with rotation was conducted. After this DNA capture step beads were pelleted on a magnetic stand and washed in ice cold Buffers C and D. Beads were re-suspended in Elution buffer and incubated at room temperature with rotation for 15 mins. After addition of Neutralisation buffer, beads were pelleted enabling enriched DNA to be separated for use in downstream applications.

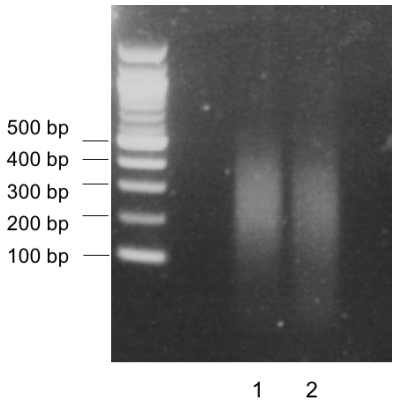


Figure 2:5 Sonication of genomic DNA for 5-hmC-DNA immunoprecipitation-sequencing/PCR. Genomic DNA was pooled from triplicate samples from day-3 differentiated mESCs exposed to atmospheric and 1% O_2 (lanes 1 and 2) and prepared for 5-hmC-DNA immunoprecipitation-sequencing/PCR by sonication into approximately 300-1000 bp fragments.

Sequencing libraries were prepared from day 3-differentiated mESCs, cultured under atmospheric O_2, using NEBNext® Ultra II™ DNA Library Prep Kit for Illumina® (New England Biolabs) following the manufacturer's protocol. Briefly, DNA fragments were end-repaired and dA tailed at the 3' end. Adaptors with a T overhang were ligated to the dA tail fragment. Fragments of 200 bp inserts were size selected using Agencourt Ampure XP Beads (Beckman Coulter), eluted in 10 mM Tris-HCl, indexed and the library PCR amplified for 12 cycles. Libraries were quantified by QPCR using NEBNext® Library Quant kit for Illumina® (New England Biolabs) and pooled at 4 nM concentration. The library pool was sequenced on the MiSeq (Illumina) single read for 50 cycles.

PCR amplification of the Tet3 promoter from atmospheric O_2 and 1% O_2 hMeDIP samples was performed using Q5® high fidelity DNA polymerase (New England Biolabs). PCR reaction mixes contained 2 µl of 5-hmC enriched DNA in 10 µl Q5 buffer, 10 µl Q5 GC Enhancer buffer, 1 µl 10 mM dNTPs, 5 µl 10 µM primers (forward and reverse), 1 µl polymerase and made with RNase/DNase-free water to give a total reaction volume of 50 µl. Primers used were as follows (given 5'-3'): forward GAGAGGGCATAGCGGACTTG and reverse GCAGACTGCAGATGAGTGGA. PCR was conducted under the following cycling conditions, according to the manufacturer's protocol: Initial denaturation (98 °C, 30 secs) and 30 cycles of denaturation (98 °C, 10 secs), annealing (64 °C, 20 secs) and extension (72 °C, 30 secs) with a final extension phase (72 °C, 2 mins). PCR products were visualised by agarose gel electrophoresis.

2.13 *TrueMethyl® genome analysis (bisulfite and oxidative-bisulfite sequencing)*

The TrueMethyl® genome analysis kit (Cambridge Epigenetix) was utilised to investigate the methylation status of specific CpGs at the single nucleotide level within the Tet3 promoter region. This assay is based on bisulfite conversion of DNA, which functions to deaminate unmethylated cytosines (C) to produce uracil (U) leaving modified 5-mC and 5-hmC bases intact. To distinguish between 5-mC and 5-hmC marks this assay kit employs the use of oxidative bisulfite chemistry. Here, the oxidative step converts C to U, 5-hmC to 5-fC and has no effect on 5-mC. After subsequent bisulfite treatment and PCR amplification, 5-mC is read as C, whereas C and 5-hmC are read as thymine (T), depicted schematically in Figure 4:21.

Genomic DNA was prepared from undifferentiated, day 3-differentiated, day 3-differentiated Tet1 KD, and day 3-differentiated 1% O_2 mESCs, as described in section

2.8.1. The size of DNA after mechanical shearing was verified by gel electrophoresis to be in the range of 300-1000 bps, see Figure 2:6. DNA was quantified using a Qubit® dsDNA HS assay kit (Thermo Scientific) and a total of 400 ng total DNA/sample was split equally between bisulfite and oxidative bisulfite conversion steps. The assay was conducted in accordance with manufacturer's instructions.

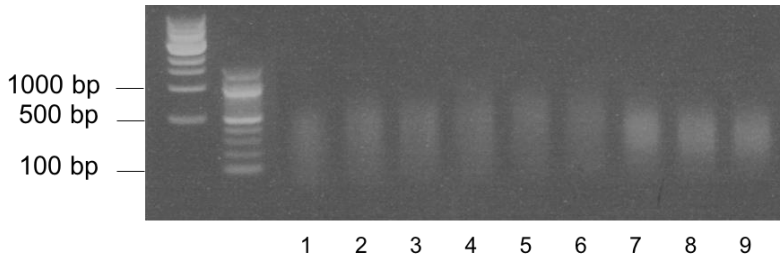


Figure 2:6 Sonication of genomic DNA for TrueMethyl® genome analysis. Genomic DNA (lanes 1-9) was prepared for TrueMethyl® genome analysis by sonication into approximately 300-1000 bp fragments.

Samples were prepared for PCR amplification in a reaction mix containing: 12.5 µl KAPA HiFi HotStart Uracil+ ReadyMix PCR kit (2X) (KAPA Biosystems), 1 µl oxidative and bisulfite treated DNA (diluted 1:10), 1.5 µl 10 µM primer and 10 µl RNase/DNase-free water. Amplification was conducted under the following cycling conditions: initial denaturation (98 °C, 45 secs), followed by 35 cycles of denaturation (98 °C, 45 secs), annealing (60 °C, 30 secs), extension (72 °C, 30 secs) and a final extension (72 °C, 1 min). Amplification products were visualised by agarose gel electrophoresis. DNA bands were excised and purified using a QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions and sequenced by Sanger Sequencing (Source BioScience) with PCR amplification primers.

Primers for oxidative and bisulfite amplification were designed using EpiDesigner primer design tool (www.epidesigner.com) with the following restrictions: amplicon size, 140 bp – 210 bp with an optimal size of 160 bp. In addition, the minimum amount of CpGs to include in the amplicon was changed from 4 to 2. Primers used were as follows (given 5'-3'): forward GATTTTTTTTAGAAGAGAAATTTGTTTAAG and reverse CAAACCAAATCAATCCTCCCTA.

2.14 *Chromatin Immunoprecipitation-PCR*

Chromatin immunoprecipitation (ChIP) is a technique used to investigate the interactions of proteins, including histone modifications, with a region of genomic DNA. Proteins are cross-linked to DNA, fragmented and immunoprecipitated with an antibody against the protein target. The DNA fragments isolated in this complex were identified by a targeted-PCR approach.

ChIP was performed using EZ-Magna ChIP™ A/G (Millipore) in accordance with manufacturer's instructions. Briefly, undifferentiated and day 3-differentiated mESCs were fixed by the addition of formaldehyde to cell culture medium at a final concentration of 1% for 10 mins at room temperature. For differentiating cells, EBs were transferred to a falcon tube and fixed under agitation. Unreacted formaldehyde was quenched with 10X glycine for 5 mins, before cells were washed twice in ice cold PBS. Cells were scraped in ice cold PBS containing protease inhibitor cocktail, collected and pelleted. Pellets were firstly lysed in cell lysis buffer for 15 mins on ice, pelleted and re-suspended in nuclear lysis buffer. DNA was sonicated to a size between 200-1000 bps, as described in section 2.8.1. For each IP reaction chromatin was mixed with immunoprecipitating antibody (see Table 2) and protein A/G magnetic beads before incubation with rotation overnight at 4 °C. An aliquot of chromatin was kept as input for PCR and a negative IgG control was included for each immunoprecipitating antibody.

Protein A/G magnetic beads were pelleted with a magnetic separator and washed in low salt buffer, high salt buffer, and LiCl wash buffer with a 5 min incubation with rotation between steps. Protein/DNA complexes were eluted, and DNA isolated by resuspension of beads in ChIP Elution Buffer with proteinase K, followed by a heat step for 2 hrs at 62 °C under agitation. Following a final heat step of 95 °C for 10 mins, DNA was purified on spin columns and eluted in 50 µl of Elution buffer C.

PCR amplification of the Tet3 promoter region was performed with Q5® high fidelity DNA polymerase. A reaction mix containing 2 µl of purified DNA template, 5 µl Q5 buffer, 5 µl Q5 GC Enhancer buffer, 0.5 µl 10 mM dNTPs, 2.5 µl 10 µM primers (forward and reverse), 0.5 µl polymerase was made up to a total reaction volume of 25 µl with RNase/DNase-free water. Primers used were as follows (given 5'-3'): forward GGGTCATCTGGTGGATCTTC and reverse GACACCGCTAGAACACAGCA. Thermocycler conditions were as follows: initial denaturation (98 °C, 30 secs) and cycles of denaturation (98 °C, 10 secs), annealing (66 °C, 20 secs) and extension (72 °C, 30 secs) with a final extension phase (72 °C, 2 mins). ChIPs for H3K27me2, H3K27me3

and H3K4me3 were amplified for 28 cycles and Tet1 for 30 cycles. Amplification products were visualised by agarose gel electrophoresis.

2.15 *Microarray analysis*

To assess the influence of Tet enzymes upon early cellular differentiation, a microarray was performed enabling simultaneous wide-spread gene expression analysis. GeneChip® technology fundamentally works by incorporating a fluorescent dye into copies of the sample and hybridising these to selected probes on an array chip, which is then scanned to produce an image file that can be analysed.

A microarray was performed from RNA isolated from day 2-differentiated P19 cells overexpressing Tet1, Tet2 and Tet3 (pcDNA™ 3.1 control was also included). The Affymetrix Mouse Transcriptome Array (MTA) 2.0 was performed and analysed by AROS Applied Biotechnology from pooled triplicate samples. Quality control analysis was performed on RNA samples prior to proceeding with the array.

2.16 *5'Rapid Amplification of cDNA ends*

To identify the transcriptional start sites of the Tet3 gene, 5'Rapid Amplification of cDNA ends (5'RACE) was performed using SMARTer® RACE 5'/3' kit (ClonTech) according to manufacturer's instructions.

In brief, 5'RACE ready cDNA was prepared from 1 μg of mRNA from day 7-differentiated mESCs. Here, first strand cDNA synthesis was primed with an oligo dt_{18} primer using SMARTScribe Reverse Transcriptase, once the 5'end of the RNA is reached, an additional few nucleotides are added to the 3'end of the first strand cDNA. This extended cDNA tag serves as a template for the SMARTer II A oligonucleotide to anneal and be reverse transcribed. This generates a cDNA copy of the original RNA with an additional tagging sequence at the end. Next, 5'RACE ready cDNA was PCR amplified using a universal primer mix (provided in the 5'RACE kit) and a gene-specific primer against mouse Tet3 (5'-3': TCTCTAGCACCATTGACCGGCGCCCCTG). A second PCR reaction using the universal short primer (provided in the 5'RACE kit)) and nested primers targeted against the potential Tet3 variants were performed (Tet3 downstream (5'-3': TGGCCCTGAGTCCATCTGAC) and Tet3 upstream (5'-3': CGCCAAGGGCACCTGAAACTGGC)). PCR products were visualised by gel

electrophoresis and extracted using the QIAquick gel extraction kit (Qiagen) according to manufacturer's instructions.

PCR products were ligated into a pGEM®-T easy vector system (Promega). This linearised vector has a single terminal T overhang at both ends enabling efficient ligation to PCR products with a terminal A overhang. TA ligations were performed in a reaction mix containing 3 μl of insert, 1 μl of pGEM®-T vector, 5 μl of 2X rapid ligation buffer and 1 μl of T4 DNA ligase. Ligations were incubated overnight and transformed as described in section 2.17.1 below. Bacteria was spread onto ampicillin (100 μg/ml) agar plates containing 100 μl of X-gal (Thermo Scientific) and incubated overnight at 37 °C. Successfully ligated products have a disrupted non-functional β-galactosidase gene that can be readily detected by white colonies in the presence of X-gal. If no insert is present, the β-galactosidase gene is active and can cleave X-gal to form blue colonies. A minimum of 6 positive clones were picked and cultured overnight at 37 °C under agitation. Plasmid DNA was purified using a QIAPrep Spin Miniprep kit (Qiagen) according to manufacturer's instructions and bi-directionally Sanger sequenced (Source BioScience) using M13 forward (GTTTTCCCAGTCACGAC) and reverse sequencing primers (CAGGAAACAGCTATGAC).

2.17 *Luciferase reporter constructs*

Luciferase reporter systems enable the instantaneous activity of cloned putative promoter sequences to be measured, thereby providing a measure of transcriptional activity. The regulatory region of a gene is cloned upstream of the luciferase gene driving its expression, thus when introduced into cells the amount of luciferase available is directly proportional to promoter activity. Once cells are lysed in the presence of luciferin substrate, the luciferase enzyme coverts luciferin into oxyluciferin and the light emitted is detected.

2.17.1 Tet3 putative promoter cloning

DNA Amplification

Tet3 putative promoter regions were amplified from 200 ng of mouse genomic DNA or from 100 ng of a Tet3 bacterial artificial chromosome (BAC) clone (RP23-331I23). Note, for the deletion series, sequences were amplified from 100 ng of the previously cloned longer constructs. Primer sequences and amplification conditions for each putative promoter fragment are shown in Table 4 and Table 5. PCR products were visualised by

gel electrophoresis, cut out and purified using the QIAquick gel extraction kit according to manufacturer's instructions. PCR products were eluted in 30 μl of RNase/DNase-free water.

Digestion and Ligation

PCR products and the pGL4.2 luciferase reporter construct (Promega) were double digested using BglII (New England Biolabs) and XhoI (New England Biolabs) in NEBuffer 3.1 (New England Biolabs) (restriction enzymes were selected for use in molecular cloning based upon their capacity to function in the same reaction buffer). Typically, 25 μl of the purified PCR product and 1 μg of vector DNA was digested with 1 μl of BglII, 1 μl of XhoI and 5 μl of 10X reaction buffer, made up to a total reaction volume of 50 μl with RNase/DNase-free water. Digestion reaction were incubated for a minimum of 1 hr at 37 °C. To prevent self-ligation of plasmid DNA, luciferase vectors were treated with 3 μl of shrimp alkaline phosphatase (SAP) (Promega) for at least 1 hr at 37 °C. Digested PCR products and vectors were purified by gel electrophoresis and extracted with the QIAquick gel extraction kit.

Ligation reactions were set up for each vector and insert to have a respective molar ratio of 1:0, 1:1, 1:3 and 1:5. Each reaction also contained 1 μl of 10X T4 DNA ligase buffer (Promega) and 1 μl of T4 DNA ligase (Promega), made up to a total of 10 μl in RNase/DNase-free water. Ligation reactions were conducted overnight at room temperature.

Transformation and colony PCR

5 μl of ligation product was mixed with 50 μl of competent JM109 bacteria cells (Promega) and incubated on ice for 1 hr. Bacteria was heat shocked (42 °C, 60 secs), placed back on ice for (5 min) and incubated with 400 μl LB under agitation (37 °C, 1 hr). 150 μl of LB-containing bacteria was spread on to ampicillin (100 μg/ml) agar plates and incubated overnight at 37 °C. Colonies were picked, and each mixed into 20 μl of RNase/DNase-free water and 4 ml of LB containing ampicillin (100 μg/ml). Bacteria-containing water samples were heated (100 °C, 5 mins), cell debris centrifuged (800 G, 5 mins) and 5 μl of lysate taken as template DNA to confirm the presence of the correct insert by PCR analysis. Each PCR reaction was performed for 25 cycles under the same optimised conditions used for initial DNA amplification (see Table 5).

Potential positive colonies were grown in the 4 ml of LB containing ampicillin (100 μg/ml), under agitation overnight at 37 °C. Plasmid DNA was purified using a QIAPrep Spin Miniprep kit according to manufacturer's instructions. Diagnostic restriction digestions on 500 ng of plasmid DNA were performed to confirm the presence of an insert (visualised by gel electrophoresis) (see Figure 2:7A). To verify whether this is the correct insert, plasmid constructs were bi-directionally Sanger sequenced) using a forward primer (Reporter vector primer 3 (Promega): CTAGCAAAATAGGCTGTCCC) upstream of the multiple cloning site on the pGL4.2 vector and the corresponding reverse cloning primer for the amplified insert. 2 ml from the starter culture of positive clones were added to 150 ml of LB containing ampicillin (100 μg/ml) and incubated overnight at 37 °C. Glycerol stocks were made by mixing the bacteria culture with sterilised glycerol in a 1:1 ratio before being stored at -80 °C. Plasmid DNA was purified using a HiSpeed® plasmid maxi purification kit according to manufacturer's instructions.

TA cloning

In some instances where the above cloning approach was unsuccessful, PCR products were firstly ligated into a pGEM®-T easy vector system. Note, PCR amplification requires the use of a non-proofreading DNA polymerase. Products amplified with the proof reading Hercules II Fusion DNA polymerase (Agilent), were subsequently PCR amplified with Taq DNA polymerase (New England Biolabs) to produce A overhangs (see Table 5 for cycling conditions). TA ligations were performed as described in section 2.16 and transformed as described above. Bacteria was spread onto X-gal containing ampicillin (100 μg/ml) agar plates and incubated overnight at 37 °C for blue/white screening (see Figure 2:7B). 3-5 white colonies were picked, cultured and plasmid DNA purified using a QIAPrep Spin Miniprep kit. The process of cloning into the pGL4.2 vector followed the steps outline above, beginning with the digestion and purification of inserts from the TA vector.

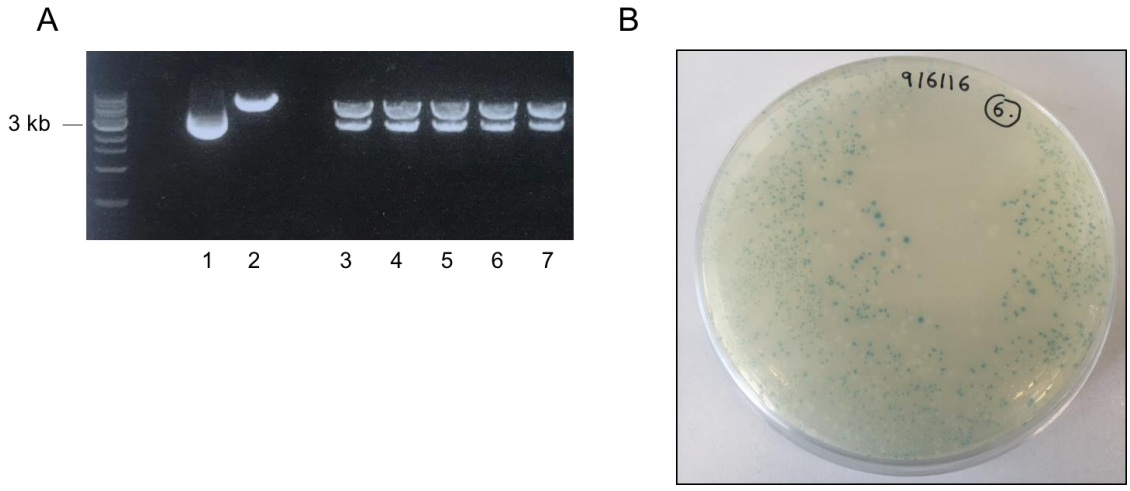


Figure 2:7 Molecular cloning. (A) Diagnostic digestion of the pGL4.2 luciferase vector to confirm the presence of the Tet3 upstream extension B insert (lanes 3-7). An empty pGL4.2 vector and its double digested linearised form is shown in lane 1 and 2 respectively. (B) The Tet3 upstream insert was ligated into the pGEM®-T easy vector system, transformed into JM109 competent cells and spread onto ampicillin agar plates containing X-gal. White colonies contain a successfully ligated insert.

2.17.2 Transfection of luciferase constructs

Cells were seeded onto 24-well tissue culture plates at a density of 60,000 cells/well for P19, SH-SY5Y and mESCs and 50,000 cells/well for H9c2 cells in 1 ml of medium. 24 hrs post plating cells were transfected with 1 µg of firefly luciferase construct DNA and 0.25 µg of a *Renilla* Luciferase construct, driven by the HSV-thymidine kinase reporter (pRL-TK) (Promega), as a co-transfected control. P19 cells were transfected with Fugene® HD, and SH-SY5Y and H9c2 cells with TurboFect™, in ratio of DNA (µg) to reagent (µl) of 1:3 and 1:2 respectively. DNA was mixed with OptiMEM, and then transfection reagent added, to give a total volume of 100 µl. Following a 15 min incubation, 100 µl of transfection mix was added to each well, culture plates gently swirled and incubated at 37 °C in a CO_2 incubator. mESCs were transfected with Lipofectamine® LTX with Plus™ Reagent (Thermo Scientific), each transfection was prepared as follows: 3 µl of LTX reagent was prepared in OptiMEM (total volume of 50 µl) and combined with a pre-prepared mix of DNA and Plus™ Reagent (in a 1:1 ratio of DNA (µg) to reagent (µl)) in 50 µl of OptiMEM. After an incubation for 5 mins, DNA-lipid complexes were added to cells (100 µl/transfection), culture plates gently swirled and incubated at 37 °C in a CO_2 incubator. For low O_2 investigations cells were exposed to 1% O_2 directly after the transfection mix was added to cells.

The activity of luciferase constructs was also investigated in differentiating P19 and mESCs. Cells were plated in T25 culture flasks at a density of 0.625 x10^6 cells/flask and 1.25 x10^6 cells/flask for P19 and mESCs respectively. 24 hrs later cells were transfected as described above, but for each transfection the amount of DNA was increased to 6 µg for the firefly luciferase reporter and to 1.5 µg for the pRL-TK control (transfection reagents were scaled accordingly). 24 hrs post-transfection cells were split equally into 3 100 mm^2 non-tissue culture grade dishes to induce differentiation and incubated at 37 °C in a CO_2 incubator, as described in sections 2.2.1 and 2.2.2.

For all assays the pGL4.2 empty vector was included to control for background luciferase activity. For each assay a minimum of 4 technical replicates for each tested construct was conducted, which for the majority of studies was performed independently a minimum of 3 times.

2.17.3 Luciferase assays

Luciferase assays were conducted using the Dual-Glo® Luciferase Assay System (Promega) according to manufacturer's instructions. In brief, 24 hrs post transfection cell medium was aspirated and 75 µl of PBS and 75 µl of the Dual-Glo® Luciferase Reagent (warmed to room temperature) was added to each well. For differentiating cells, aggregates were gently centrifuged (800 G), medium aspirated and re-suspended in the PBS and Dual-Glo® substrate mix. Following a 10 min incubation, under agitation, 140 µl of lysate was transferred to a single well of a 96-well flat bottomed white opaque plate (Corning) and firefly luminescence signal read on a Mithras LB 940 luminometer (Berthold Technologies). 70 µl of Dual-Glo® Stop & Glo® Reagent was added to each well for 10 mins under agitation and the *Renilla* luciferase signal read. The activity of the firefly luciferase constructs were normalised to the *Renilla* signal and expressed relative to the pGL4.2 empty vector construct.

2.18 *Animal Husbandry*

Mice were housed in a standard temperature and humidity controlled animal facility with a 12 hr light/dark cycle (daylight from 7 am to 7 pm). In accordance with Housing and Care of Animals Used in Scientific Procedures Code of Practice, mice were allowed free access to water and standard laboratory chow. All experimental mice used were on a C57BL/6 background. Animals were killed according to Schedule 1 methods and tissues harvested as required.

2.19 *Data analysis*

Each experiment was performed to include at least a minimum of a technical triplicate, and where possible was biologically repeated 3 times. Data was statistically analysed and made into figures using GraphPad Prism 7 software. Data containing 2 groups was analysed by an un-paired student t-test. Data containing more than 2 groups where analysed by a one-way ANOVA with post-hoc analysis. A Dunnett's test was used when comparisons were made to a control group, whereas a Tukey test was used to compare all sample groups to each other. In some instances where there were 2f changing variables a two-way ANOVA was utilised. In these circumstances the Sidak multiple comparison test was also used to enable comparisons between selected groups of data. Where possible, data was expressed at mean ± standard error of the mean (SEM).

Next-generation sequence analysis was performed using BaseSpace software (Illumina) and Galaxy software (https://usegalaxy.org) against the mouse mm10 reference genome. Data was visualised using the Integrative Genomic Viewer (https://software.broadinstitute.org/software/igv/).

Reagents	Resolving gel 7.5% (µl)	Resolving gel 15% (µl)	Stacking gel 7.5% (µl)	Stacking gel 15% (µl)
30% Acrylamide	2500	5000	680	680
1.5 M Tris-HCl pH 8.8	2500	2500	-	-
1 M Tris-HCl pH 6.8	-	-	500	500
10% Sodium Dodecyl Sulfate (SDS)	100	100	40	40
10% Ammonium persulfate (APS)	100	100	40	40
Tetramethylethylenediamine (TEMED)	10	10	4	4
Water	4790	2290	2736	2736

Table 1: Western blot gel casting reagents

Antibody	Dilution
5-mC (Actif Motif, 39649)	**DB:** 1:1000 in Odyssey® blocking buffer (PBS)
5-hmC (Actif Motif, 39770)	**DB:** 1:10000 in 10% milk PBS/T
TET1(Millipore, 09-872)	**WB**: 1:1000 in 5% milk TBS/T **ChIP**: 1:50
TET2 (AbCam, ab94580)	**WB**: 1:1000 in 5% milk TBS/T
TET3 (AbCam, ab139311)	**WB**: 1:500 in 5% milk TBS/T
H3K27me2 (Cell Signalling, 9728)	**WB**: 1:1000 in 10% milk TBS/T
H3K27me3 (Actif Motif, 39155)	**WB**: 1:1000 in 10% milk TBS/T **ChIP**: 1:50
H3K4me3 (Cell Signalling, 9751)	**WB**: 1:1000 in 5% BSA TBS/T **ChIP**: 1:50
HIF (Thermo Scientific, 16511)	**WB**: 1:1000 in 10% milk PBS/T
FLAG (Sigma, F3165)	**WB**: 1:1000 in 10% milk TBS/T
β-ACTIN (Sigma, A5316)	**WB**: 1:5000 in 10% milk TBS/T
α-TUBILIN (Sigma, T5168)	**WB**: 1:10000 in 10% milk TBS/T

Table 2: List of primary antibodies. DB: dot-blot, WB: western-blot, ChIP: chromatin immunoprecipitation, TBS/T: Tris-buffered salt solution containing 0.1% tween-20, PBS/T: Phosphate buffered saline containing 0.1% tween-20, BSA: bovine serum albumin

Gene	Forward primer 5'-3'	Reverse primer 5'-3'
β-III tubulin	TTTTCGTCTCTAGCCGCGTG	GATGACCTCCCAGAACTTGGC
Brachyury	GTATTCCCAATGGGGGTGGCT	CGAGCCTCCAAACTGAGGG
Canx	TTCCAGACCCTGSTGCAGA	TCCCATTCTCCGTCCATATC
Dcx	ACGACCAAGACGCAAATGGA	CTTCTGCTTCCGCAGACTTC
Dnmt3b	CCCTCCCCCATCCATAGT	TCTGCTGTCTCCCTTCATTGT
Ezh2	CAGGCTGGGGCATCTTTATC	ACGAATTTTGTTGCCCTTTC
Fabp7	AACCAGCATAGATGACAGAAACTG	ACTTCTGCACATGAATGAGCTT
Foxa2	CATCCGACTGGAGCAGCTA	TGTGTTCATGCCATTCATCC
Foxj3	TGAACAGTGTTGGAAGTGTACATAGT	CTGCTGCTGCTGAGGTGTT
Gapdh	GGGTTCCTATAAATACGGACTGC	CCATTTTGTCTACGGGACGA
Gata2	CTCCAGCTTCACCCCTAAGC	ACCACAGTTGACACACTCCC
Gata4	GGAAGACACCCCAATCTCG	CATGGCCCCACAATTGAC
Gdf3	GGGTGTTCGTGGGAACCT	CCATCTTGGAAAGGTTTCTGT
Glut1	GGCGGGACACGCATAGTT	ACATGAGGCGGCCCGT
Glut3	GACCCGAGGAACACTTGCTG	CGAACACCAGAGATGGGGTC
Gulo	AAGCCGGTATCCTCCTGACT	CCACCAACCGTCACATCAGA
Hand1	CGGAAAAGGGAGTTGCCTCA	GGTGCGCCCTTTAATCCTCT
Hes5	CCCAAGGAGAAAAACCGACT	TGCTCTATGCTGCTGTTGATG
Hnf4a	ACACCACCCTGGAGTTTGAA	GCCCAGGCTGTTGGATGAAT
Lefty1	CAGCTCGATCAACCGCCAGT	GGCTGGCATGGCTGTGTT
Map2	TTCCCTCGTTTCTTCGGTCG	GCACACAGCACAGCCTGG
Mesp1	CGCCTGCCTACCCTAGAC	TGAAGAGCGGAGATGAGGGA
Myh7	AGCAGCAGTTGGATGAGCGACT	CCAGCTCCTCGATGCGTGCC
Nanog	AAGGATGAAGTGCAAGCGGT	GGTGCTGAGCCCTTCTGAAT
Ncam1	AAGTACAGAGCGCTCGCC	AGGGACTTGAGCATGACGTG
Nestin	AGGCGCTGGAACAGAGATTG	CACAGCCAGCTGGAACTTTT
NeuroD	CGAGGCTCCAGGGTTATGAG	TTGGTCATGTTTCCACTTCCTGT
Nppa	CAACACAGATCTGATGGATTTCA	CCTCATCTTCTACCGGCATC
Oct4	GTTGGAGAAGGTGGAACCAA	CTCCTTCTGCAGGGCTTTC
Snail1	GTCTGCACGACCTGTGGAAA	GGTCAGCAAAAGCACGGTTG
Sox2	GCACATGAACGGCTGGAGCAACG	TGCTGCGAGTAGGACATGCTGTAGG
Sox7	AGATGCTGGGAAAGTCATGG	AGAGGGAGCTGAGGAGGAAG
Svct1	CGGGTGCCTTGCTTTCATAC	CTGGCTGACGTCTCACTGTT
Svct2	GTGCGATCGGCGGGAC	ACCAGGGACACTGGAGTCTT
Tbx5	AGCACTTCTCCGCTCATTTCAC	TGAGGTCTGGTGCTGAAACATC
Tet1	GAGCCTGTTCCTCGATGTGG	CAAACCCACCTGAGGCTGTT
Tet1 h	CCGAATCAAGCCGAAGAATA	CCTGGAGATGCCTCTTTCAC
Tet2	TGTTGTTGTCAGGGTGAGAATC	TCTTGCTTCTGGCAAACTTACA
Tet3	CCGGATTGAGAAGGTCATCTAC	AAGATAACAATCACGGCGTTCT
Tet3 down	GGCCGATGCAGTAGTTGGAG	TTGACAGTCGCCCCTTGT
Tet3 h	CCATTGCAAAGTGGGTGA	CGCACCAGGCAGAGTAGC
Tet3 up	TTGTACCAGCTGCACCAATC	CGTTCCTTCACGAGCATTTA

Table 3: List of QPCR primers. h:human, Tet3 up: Tet3 upstream, Tet3 down: Tet3 downstream

Promoter Fragment	Construct	Forward primer 5'-3'	Reverse primer 5'-3'	Fragment size (kb)
Tet3 downstream putative promoter	Tet3 downstream	*GGAAG*CTCGAGGGCGAAGGATATGGTCTTCA	*GGAAG*AGATCTGCCAGAGGCTACTGCTGAGT	1.93
	Tet3 downstream deletion 1	*GGAAG*CTCGAGGCCTGGCTGCTGGGCAGGTC	*GGAAG*AGATCTGCCAGAGGCTACTGCTGAGT	1.55
	Tet3 downstream deletion 2	*GGAAG*CTCGAGGGCGAAGGATATGGTCTTCA	*GGAAG*AGATCTACACACACAAACACGCAGGG	0.94
	Tet3 downstream deletion 3	*GGAAG*CTCGAGGGCGAAGGATATGGTCTTCA	*GGAAG*AGATCTCTCCCTTCCAGGCCCAGCGC	0.61
Tet3 upstream putative promoter	Tet3 upstream	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTCAAAGGGGCAAAGTTTCCCA	1.50
	Tet3 upstream deletion A	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTTTCCTGCCACTCCCTGCGAG	0.92
	Tet3 upstream extension A	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTATGATGGAGAAGGCTACGCT	2.36
	Tet3 upstream extension B	*CCTT*CTCGAGGACTGTTGATATTTCCATAC	*GCTTC*AGATCTCAAAGGGGCAAAGTTTCCCA	2.95
	Tet3 upstream deletion 1A	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTCCTCAGAGGGAGGTACATAG	2.20
	Tet3 upstream deletion 2A	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTCTCGCTCATTTGGAGTGGAC	1.97
	Tet3 upstream deletion 3A	*CCTT*CTCGAGGTTAACCCTCCCTACCCACC	*GCTTC*AGATCTGGGTCGCTGGCTACCCTG	1.65

Table 4: A list of cloning primers for amplification of the Tet3 putative promoter regions. Cloning primers contain a leader sequence, shown in italics, a restriction enzyme site, underlined, and a complementary sequence to the region of interest.

Promoter Fragment	Construct	PCR reaction Mix	PCR amplification conditions
Tet3 downstream putative promoter	Tet3 downstream	200 ng mouse genomic DNA, 0.5 μl 10 mM dNTPs (Promega), 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer (Agilent), 0.4 μl DMSO and 1 μl Hercules II Fusion DNA polymerase (Agilent), made up to a 50 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (65 °C, 20 secs), extension (72 °C, 2.5 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 downstream deletion 1	100 ng Tet3 downstream construct, 12.5 μl 2X REDTaq (Sigma), 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, made up to 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (94 °C, 20 secs), annealing (55 °C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 downstream deletion 2	100 ng Tet3 downstream construct, 12.5 μl 2X REDTaq, 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, made up to 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (94 °C, 20 secs), annealing (55 °C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 downstream deletion 3	100 ng Tet3 downstream construct, 12.5 μl 2X REDTaq, 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, made up to 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (94 °C, 20 secs), annealing (55 °C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)
Tet3 upstream putative promoter	Tet3 upstream	200 ng mouse genomic DNA, 0.5 μl 10 mM dNTPs, 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 0.8 μl DMSO and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (48°C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)

	Tet3 upstream deletion A	100 ng Tet3 upstream construct, 12.5 μl 2X REDTaq, 1.25 μl 10 μM forward primer, 1.25 μl 10 μM reverse primer, made up to 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (94 °C, 20 secs), annealing (62 °C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 upstream extension A	200 ng Tet3 BAC clone, 2.5 μl 10 mM dNTPs, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 10 μl betaine and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water For TA cloning: 4 μl of Hercules II amplified PCR product, 0.5 μl 10 mM dNTPs, 0.5 μl 10 μM forward primer, 0.5 μl 10 μM reverse primer, 2.5 μl of 10X standard Taq reaction buffer (New England Biolabs), 1 μl Taq polymerase (New England Biolabs), made up to a 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 20 secs), extension (72 °C, 3 mins), and a final extension phase (72 °C, 3 mins) For TA cloning: Initial denaturation (95 °C, 30 secs), followed by 10 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 30 secs), extension (68 °C, 3 mins), and a final extension phase (68 °C, 5 mins)
	Tet3 upstream extension B	200 ng Tet3 BAC clone, 2.5 μl 10 mM dNTPs, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 10 μl betaine and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water For TA cloning: 4 μl of Hercules II amplified PCR product, 0.5 μl 10 mM dNTPs, 0.5 μl 10 μM forward primer, 0.5 μl 10 μM reverse primer, 2.5 μl of 10X standard Taq reaction buffer, 1 μl Taq polymerase, made up to a 25 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 20 secs), extension (72 °C, 3 mins), and a final extension phase (72 °C, 3 mins) For TA cloning: Initial denaturation (95 °C, 30 secs), followed by 10 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 30 secs), extension (68 °C, 3 mins), and a final extension phase (68 °C, 5 mins)

	Tet3 upstream deletion 1A	100 ng Tet3 upstream extension A, 2.5 μl 10 mM dNTPs, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 10 μl betaine and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 20 secs), extension (72 °C, 3 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 upstream deletion 2A	100 ng Tet3 upstream extension A, 2.5 μl 10 mM dNTPs, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 10 μl betaine and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 20 secs), extension (72 °C, 2.5 mins), and a final extension phase (72 °C, 3 mins)
	Tet3 upstream deletion 3A	100 ng Tet3 upstream extension A, 2.5 μl 10 mM dNTPs, 2.5 μl 10 μM forward primer, 2.5 μl 10 μM reverse primer, 10 μl 5X Hercules II reaction buffer, 10 μl betaine and 1 μl Hercules II Fusion DNA polymerase, made up to a 50 μl volume with RNase/DNase-free water	Initial denaturation (95 °C, 2 mins), followed by 30 cycles of denaturation (95 °C, 20 secs), annealing (52°C, 20 secs), extension (72 °C, 2 mins), and a final extension phase (72 °C, 3 mins)

Table 5: PCR reaction and amplification conditions for Tet3 promoter cloning

Chapter 3: Results 1; *The Role of O_2 and Tet Enzymes in Cellular Differentiation*

3.1 *Introduction*

The differentiation of ESCs *in vitro* into 3-dimensional cell aggregates, termed EBs recapitulates the early stages of embryonic development. A commonly used and robust way to induce EB formation is to plate ESCs in suspension on a non-adherent surface. Cells self-assemble themselves together *via* cell-cell adhesions[298,299], proliferate and spontaneously differentiate[300]. The formation of the spheroid structure of EBs in part mimics the early embryo and provides a model to study the mechanisms controlling the spacio-temporal patterns of gene expression. Specifically, the way in which cells become asymmetrically or unequally distributed is poorly understood, but the significance of morphogen gradients has long been suggested[301]. O_2 is an important regulator of embryogenesis and is considered a developmental morphogen[69], yet fundamentally how this is sensed to enable developmental patterning remains to be fully elucidated.

3.1.1 Tet activity and O_2

Tet enzymes as described in detail previously (see Chapter 1: General Introduction) belong to members of the 2-OGDO family. The activity of this group of enzymes is absolutely dependent upon its substrates, 2-OG and O_2[302]. Therefore, alterations in the available levels of these substrates have the capacity to alter the activity of these enzymes and potentially elicit a biological response. Perhaps the most studied example of this is the dioxygenase component of HIFs, PHDs, whose hydroxylation activity is inhibited in low [O_2], resulting in the stabilisation and activation of HIF to enable its downstream signalling. Unlike PHDs, the ability of Tets to act as cellular O_2 sensors remain to be elucidated.

To be an O_2 sensor an enzyme must have a high K_m (the concentration of substrate required for an enzyme to achieve half its maximal response (V_{max})) and thus a low affinity for O_2, so that its activity can be varied over physiological changes in [O_2]. The K_m of Tet1 and Tet2 has been calculated to approximately 30 μM[303] (equating to ~3% O_2). This suggests Tet1 and 2 have a high affinity for O_2, which in some instances in the developing embryo may allow them to be catalytically active, but considering the physiological [O_2] may be as low as 1% O_2[70,71], these enzymes could demonstrate spatial inhibition. However, it should be considered that the calculated K_m for PHDs were between 230 and 250 μM, which exceeds the [O_2] in air[304] (200 μM, calculated in air-saturated aqueous buffer at 37 °C[305]). These K_m values are determined from *in vitro* enzymatic measurements *via* detection of $^{14}CO_2$ release from 2-oxo-[1-^{14}C]-glutarate. A caveat underpins this methodology in that the enzyme-substrate is most likely to bind

O_2, thus this order of binding implies the substrate bound to the enzyme could affect the K_m of O_2[306,307]. This is in agreement with crystallography data on FIH-1, whereby the preformed enzyme-Fe^{2+}-2-OG substrate complex was found to bind O_2[308], thereby suggesting K_m values should be considered with caution.

A recent study, which to date provides the most in-depth insight into O_2 and Tet activity, revealed exposure of mESCs to 0.5% O_2, compared to atmospheric controls, reduced 5-hmC levels in mESCs[309]. Further, 5-hmC loss was also confirmed from hypoxic tumours, and was subsequently found to be associated with hypermethylated gene promoters[309]. This study also attempted K_m calculations through measuring the conversion of 5-mC to 5-hmC of recombinant purified Tet1 and Tet2, exposed to decreasing [O_2]. K_m values of 0.31% and 0.53% O_2 was determined for Tet1 and Tet2 respectively[309], which highlights a high affinity for O_2, which may suggest an unlikelihood of such enzymes to be physiological O_2 sensors. However, the catalytic activity of all 3 paralogues in response to O_2, in a cellular model, remain unexplored, and furthermore, the concept of these enzymes as O_2 sensors in embryogenesis remain to be elucidated.

3.1.2 Cellular metabolites and Tet activity

Cancer studies have revealed mutations in 2 TCA cycle genes, fumarate hydratase (FH) and succinate dehydrogenase (SDH)[310,311], result in accumulation of fumarate and succinate respectively. These substrates have been shown to be competitive inhibitors of 2-OGDO enzymes, which include the Tets[312] (Schematically shown in Figure 3:1). Specifically, these metabolites have been shown to achieve potent IC_{50} values between 400-500 μM for Tet1 and Tet2[303] and have been shown to have a functional affect by altering genome-wide DNA methylation profiles[312].

Succinate and fumarate also have been shown to inhibit PHDs, by stabilising HIF and activating subsequent downstream signalling, thereby acting as a 'pseudo-hypoxic' response, enhancing angiogenesis and anaerobic metabolism in tumours[313,314,315,316]. The ability of these TCA metabolites to induce a 'hypoxic response' may also arise from changes in cellular energetics in response to O_2. The metabolic state of the cell is dependent upon [O_2], which is required for ATP production *via* the OxPhos pathway. To sustain cell viability if [O_2] decreases, a metabolic switch from OxPhos to anaerobic glycolysis occurs, to maintain, albeit inefficiently, ATP generation[317,318]. This would thus result in a decrease in TCA metabolites, reducing levels of the 2-OGDO co-substrate 2-OG, and therefore potentially reducing enzymatic activity. Furthermore, levels of succinate in the mitochondrial matrix has been demonstrated to increase from 0.5 mM

to 6 mM under ischemic conditions *in vivo*[319,320]. Therefore, a low [O_2] may induce indirect regulation of Tet activity through altering the metabolic signature of the cell and hence inducing changes in metabolic inhibitors of Tet enzymes.

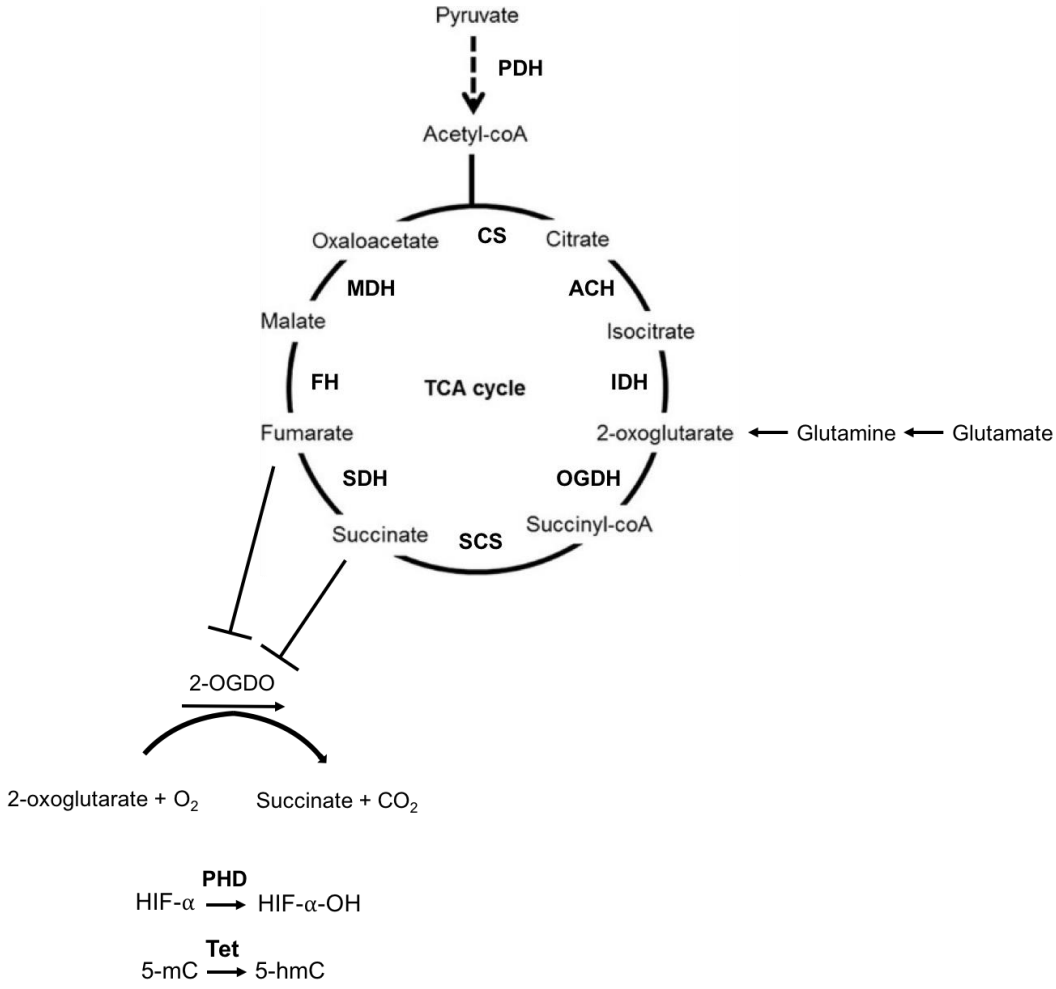


Figure 3:1 TCA cycle substrates inhibit 2-OGDOs. Pyruvate, the product of glycolysis, is converted to acetyl coA, which enters the TCA cycle and is converted to high energy intermediates. This energy is harvested through electron transport by the oxidative phosphorylation (OxPhos) pathway where it is converted to ATP. Fumarate and succinate are inhibitors of 2-OGDOs and have specifically been shown to inhibit PHDs, resulting in the stabilisation of HIF and Tet enzymes, resulting in reduced hydroxylation and thus DNA hypermethylation.

PDH: pyruvate dehydrogenase, CS: citrate synthase, ACH: aconitate hydratase, IDH: isocitrate dehydrogenase, OGDH: oxoglutarate dehydrogenase, SCS: succinyl-coA synthetase, SDH: succinate dehydrogenase, FH: fumarate hydratase, MDH: malate dehydrogenase.

3.1.3 Distinct roles of Tet1, Tet2 and Tet3 in cellular differentiation

Tet proteins are functionally intrinsic to embryogenesis, detailed within Chapter 1: General Introduction. However, the precise role of each individual paralogues in cellular differentiation remains largely unknown, in part due to the functional redundancy between Tet paralogues and the variety of cell culture conditions utilised in cellular differentiation models. The general accepted dogma (in part owing to their expression profiles) is that Tet1 and Tet2 are involved in the pluripotency and self-renewal processes, whereas Tet3 is implicated in differentiation events[252,321,322].

The limited studies that have been reported have attempted to delineate between the functional roles of each paralogue. For example, in mESCs Tet1 and Tet2 have been shown to regulate 5-hmC at different genomic locations, Tet1 is responsible for 5-hmC production at promoter regions, whereas Tet2 regulates 5-hmC deposition at gene bodies[323]. This may potentially account for the functional differences determined from the loss-of-function studies in mESCs of Tet1 and Tet2, whereby differential skewing towards the endoderm/mesoderm/trophoblast lineage and towards the neuroectoderm respectively, was reported[252]. In agreement with this, it was independently shown that Tet1 depleted mESCs also skewed differentiation towards the trophectoderm[251] and Tet2 silenced hESCs altered differential patterning towards the neuroectodermal lineage[324].

Tet3 has been implicated in cellular specification, and specifically induction of neurogenesis[325], a topic that is covered in Chapter 5: Results 3. However, it is important to highlight here that the silencing of Tet3 in mESC and in neural progenitor cells (NPCs), resulted in impairment of neuronal differentiation and the formation of terminal differentiated neurones respectively[321]. More recently, it has also been shown that when induced to differentiate along a neural lineage, Tet3-deficient mESCs could skew differentiation away from the neuroectoderm and towards the mesodermal lineage[322]. In addition to neural specification, Tet3 is also linked to differentiation events including erythroid differentiation[326] and eye formation[327]. Thus, it can be inferred Tet3 has specific roles for regulating cell fate decisions during differentiation.

3.2 *Aims*

To investigate the effects of low [O_2] upon cellular differentiation and determine whether any identified gene expression changes may in part be regulated by Tet enzymes. To evaluate this, the aim is to determine the ability of Tet enzymes to function as O_2 sensors, and consider their potential to differentially regulate developmental genes during cellular differentiation.

3.3 *Results*

3.3.1 P19 cells form all three germ layers when induced to differentiated into embryoid body-like cell aggregates

Murine P19 embryonal carcinoma cells, derived from teratocarcinoma-induced mice, are pluripotent cells that can differentiate into cell types of all 3 germ layers and thus resemble early embryonic cells[328,329]. These cells are easily expanded and maintained in an undifferentiated state, but can also be efficiently induced to differentiate, thereby providing a model to begin to elucidate the effect of O_2-dependent signalling upon cellular differentiation.

To characterise the patterns of lineage marker expression, P19 cells were induced to differentiate unbiasedly into EB-like aggregates and were harvested at days 3,7 and 11 for QPCR analysis. Firstly, the expression of key pluripotent genes, Oct4[330,331] and Nanog[332–334] were confirmed in undifferentiated P19 cells (Figure 3:2). As might be expected, expression of Oct4 was decreased significantly at days 7 and 11, compared to undifferentiated cells. However, no significant change in Nanog expression was observed, although expression appeared to be decreased in a time-dependent manner from days 3 to 11.

The expression of key germ layer markers were also assessed over the defined differentiation time course (Figure 3:3). Endodermal markers, Foxa2[335] and Sox7[336,337], displayed distinct expression profiles. Foxa2 was significantly elevated at day 3 to approximately 6-fold, relative to undifferentiated cells, and then decreased back to comparable expression levels of the undifferentiated state by day 11. In comparison Sox7 showed a burst of expression at day 7. Early mesodermal markers, Brachyury[338], Hand1[339] and Mesp1[340] were all significantly elevated at day 3, with Brachyury showing the most extensive induction of around 90-fold. Expression was subsequently reduced to levels comparative of undifferentiated cells at day 7 for Brachyury and Mesp1 and day 11 for Hand1. The expression of cardiac progenitor markers Tbx5[341] and Gata4[342] were significantly elevated at later time points, day 11 and 7 respectively. Ectodermal markers, specifically β-III tubulin[343] and Ncam1[344] (despite not reaching significance) were also elevated at later time points (days 11 and 7 respectively). Other ectodermal markers, Nestin[345], Foxj3[346] and Map2[347], showed little change over the entirety of the differentiation time course, and only Foxj3 displayed significantly reduced expression at day 3.

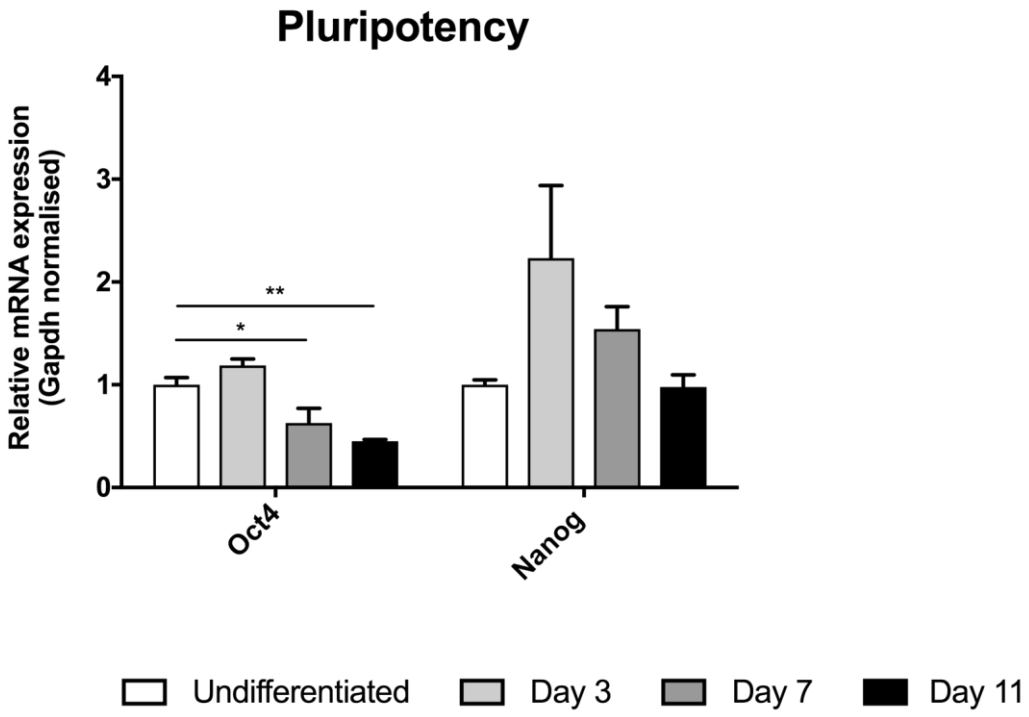


Figure 3:2 Pluripotency marker mRNA expression during a time course of P19 cell differentiation. P19 cells were differentiated over an 11 day time course and mRNA expression of Oct4 and Nanog assessed by QPCR analysis. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. * $p<0.05$, ** $p<0.01$

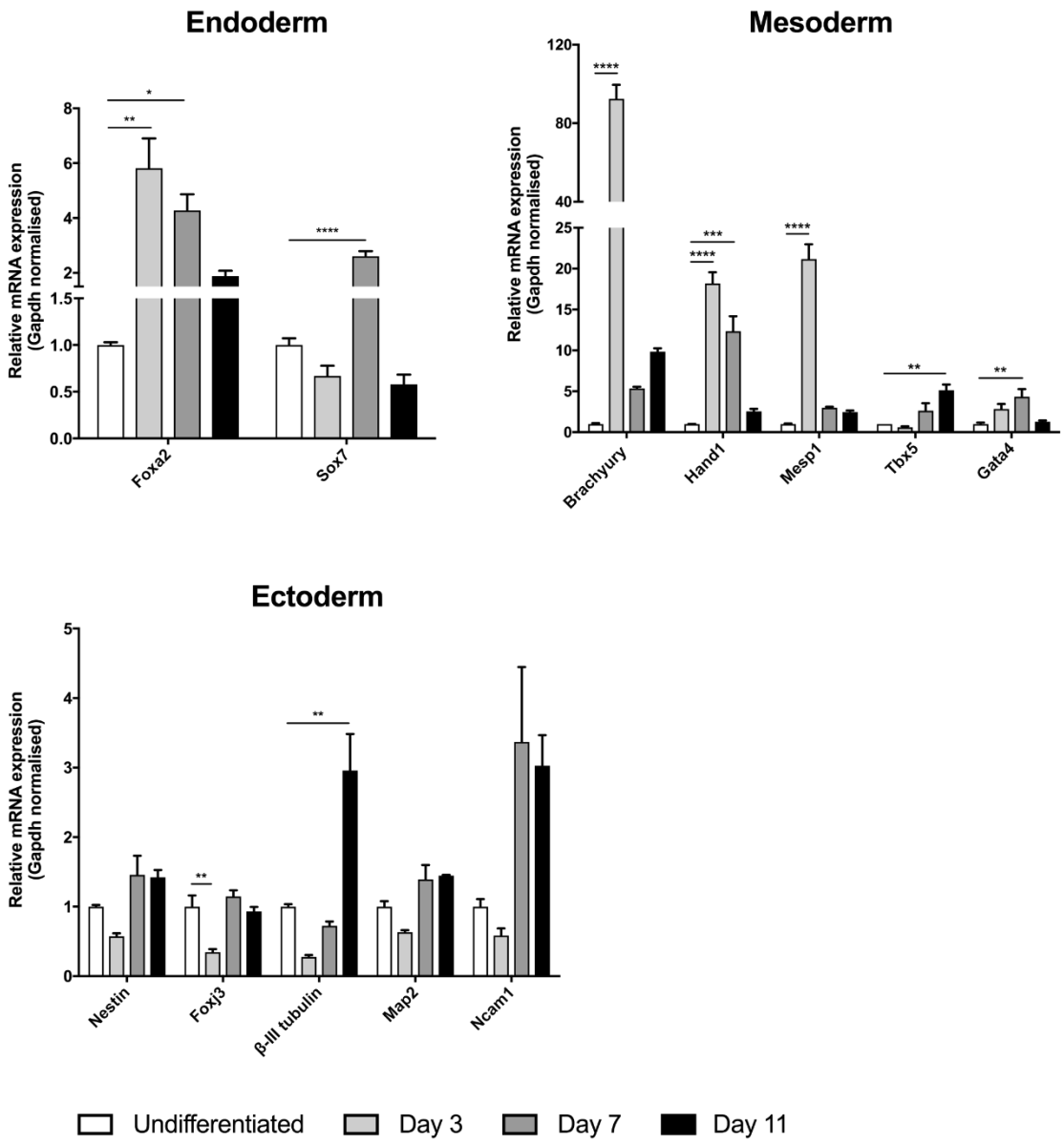


Figure 3:3 mRNA expression of genes representative of all three germ layers during a time course of P19 cell differentiation. P19 cells were differentiated over an 11 day time course and mRNA expression of key lineage markers were assessed by QPCR. Data are expressed as the mean ± SEM and for each gene analysed by a one-way ANOVA with Dunnett's post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

3.3.2 Chemically-mimicked hypoxia modulates gene expression of pluripotent and lineage markers in differentiating multipotent P19 cells

To investigate the effects of O_2-dependent signalling pathways upon cellular differentiation, P19 cells were induced to differentiate into EB-like cell aggregates in the presence of the non-specific 2-OGDO inhibitor DMOG[348]. Activation of low [O_2] mediated cellular signalling *via* DMOG was confirmed by induction of HIF-1α protein expression (Figure 3:4).

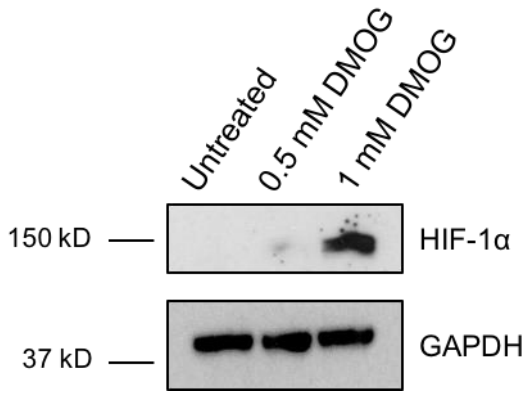


Figure 3:4 HIF-1α protein is induced by 1 mM Dimethyloxalylglycine treatment. Undifferentiated P19 cells were treated with 0.5 and 1 mM DMOG for 16 hrs, harvested and probed for HIF-1α.

In separate studies, cells were either differentiated for 4 days in the presence of DMOG or pre-treated with DMOG for 4 days and differentiated further till days 7 and 11. Cells were harvested for QPCR analysis and gene expression of pluripotent and lineage markers was assessed. These studies aimed to inform whether a constant mimicked hypoxic environment can alter the differentiation profile and/or whether an early hypoxic stimulus can affect cell specification at later time points.

To identify whether DMOG induces or inhibits differentiation of P19 cells, expression of pluripotency markers was assessed (Figure 3:5). DMOG treatment significantly decreased Oct4, while trended to increase Nanog expression after 4 days. However, pre-treatment with DMOG augmented expression of both genes at day 7, but this was not significantly maintained by day 11. This suggests that an initial 'hypoxic' environment for 4 days with DMOG acted to maintain pluripotency until day 7 and therefore delayed the onset of differentiation. The expression of the endodermal markers (Figure 3:6),

Sox7 and Foxa2 were respectively found to be decreased and unaffected after DMOG treatment at day 4. This effect was sustained for Sox7 at day 7 after pre-treatment of DMOG. In comparison, Foxa2 expression was lost at days 7 and 11 after DMOG pre-treatment. Mesodermal markers displayed a more inconsistent profile (Figure 3:7), Hand1 was induced by 14-fold while Brachyury (despite failing to reach significance) trended to increase after 4 days of DMOG treatment. However, the expression of Gata4 and Mesp1 was ablated at this time point. Pre-treatment with DMOG maintained this effect upon Brachyury and Gata4 expression till days 7 and 11 respectively, although it had an opposing effect on Hand1 expression at day 7. Ectodermal markers (Figure 3:8), Ncam1 and Nestin were notably decreased after 4 days of DMOG treatment, an effect maintained for Nestin after pre-treatment at days 7 and 11. Similar to Nestin, β-III tubulin expression was increased in control cells between days 7 and 11, but markedly reduced after DMOG pre-treatment. Foxj3 was unaffected by DMOG directly and after pre-treatment at both tested time points.

Taken together, these data demonstrate the ability of DMOG to skew differentiation away from endoderm and ectodermal lineages, and disrupt mesodermal formation. Therefore, this is consistent with a role for O_2-sensitive cell signalling in the orchestration of embryonic events, and highlights the potential of O_2 to act as a morphogenic signal.

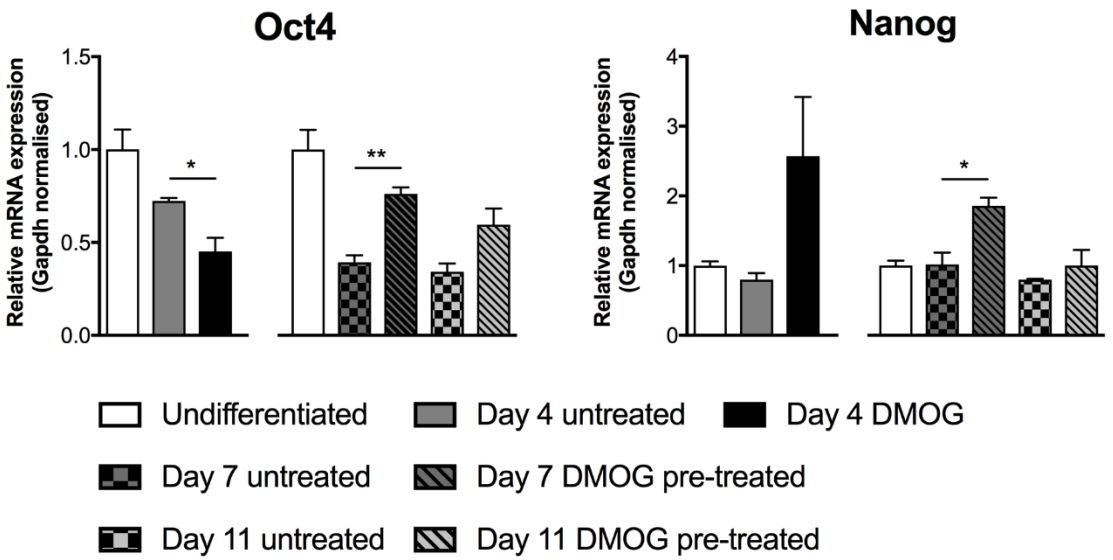


Figure 3:5 Chemically-mimicked hypoxia by dimethyloxalylglycine treatment and pretreatment alters pluripotency marker expression. In 2 independent studies P19 cells were induced to differentiate in the presence of 1 mM DMOG for 4 days, or were pre-treated with DMOG for 4 days before allowed to differentiate for a further 11 days. Pluripotency marker expression was analysed by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by an unpaired t-test between DMOG treatments at each time point. * $p<0.05$, ** $p<0.01$

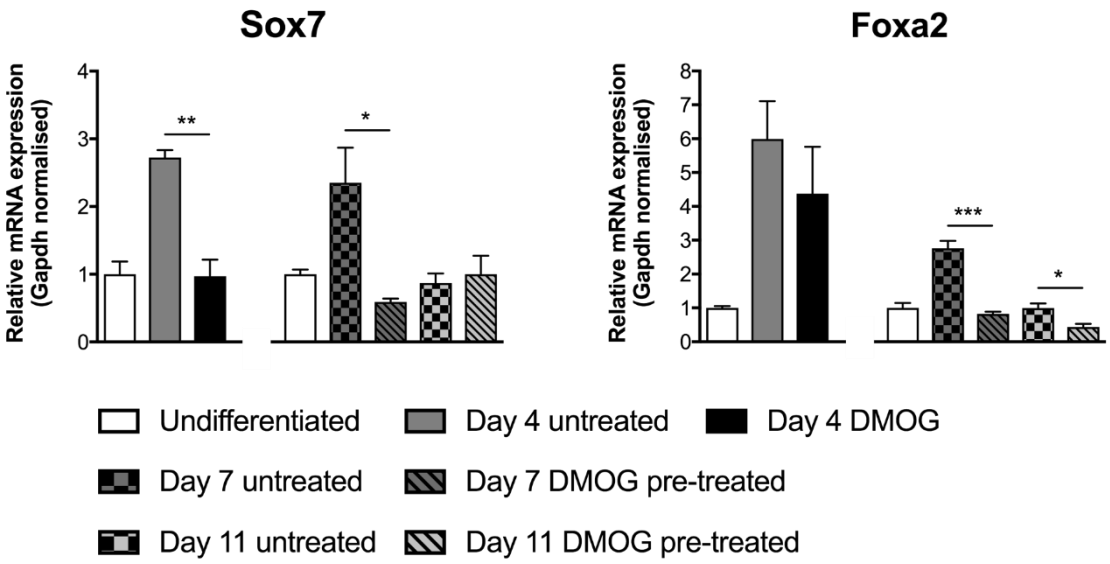


Figure 3:6 Chemically-mimicked hypoxia by dimethyloxalylglycine treatment and pretreatment decreases endodermal marker expression. In 2 independent studies P19 cells were induced to differentiate in the presence of 1 mM DMOG for 4 days or were pre-treated with DMOG for 4 days before allowed to differentiate for a further 11 days. Endodermal marker expression was analysed by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by an unpaired t-test between DMOG treatments at each time point. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

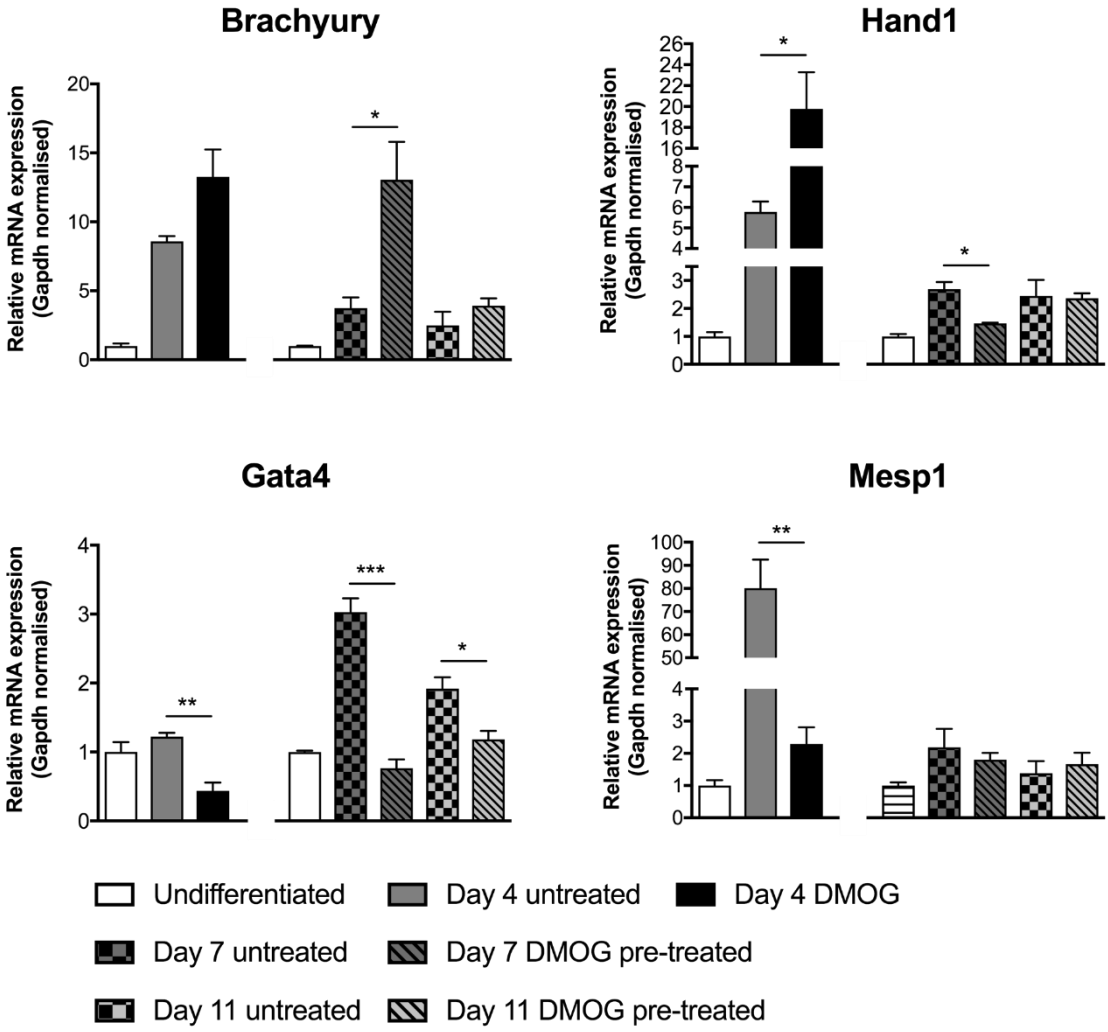


Figure 3:7 Chemically-mimicked hypoxia by dimethyloxalylglycine treatment and pretreatment regulates mesodermal marker expression. In 2 independent studies P19 cells were induced to differentiate in the presence of 1 mM DMOG for 4 days or were pre-treated with DMOG for 4 days before allowed to differentiate for a further 11 days. Mesodermall marker expression was analysed by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by an unpaired t-test between DMOG treatments at each time point. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

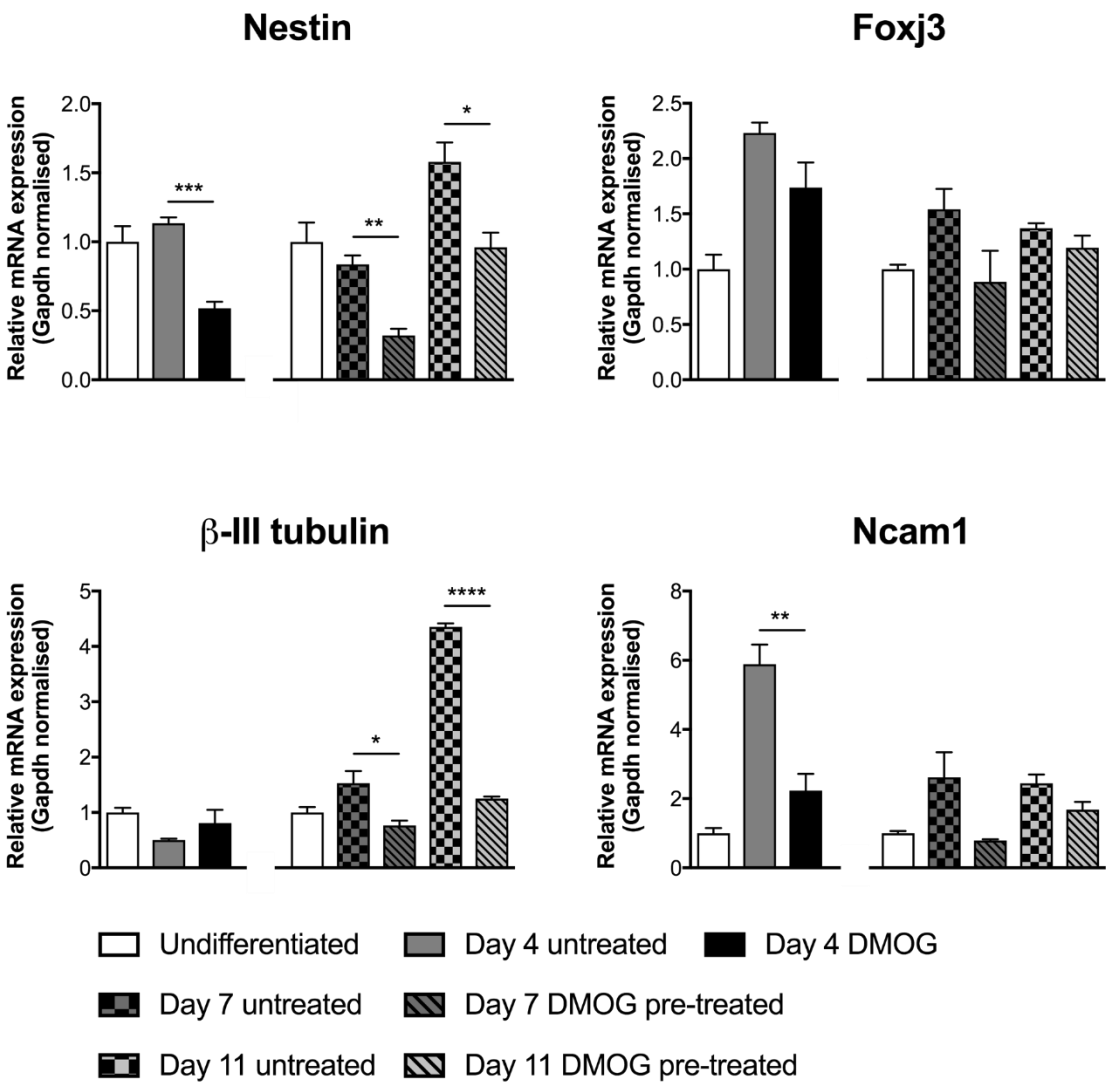


Figure 3:8 Chemically-mimicked hypoxia by dimethyloxalylglycine treatment and pretreatment decreases ectodermal marker expression. In 2 independent studies P19 cells were induced to differentiate in the presence of 1 mM DMOG for 4 days or were pre-treated with DMOG for 4 days before allowed to differentiate for a further 11 days. Ectodermal marker expression was analysed by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by an unpaired t-test between DMOG treatments at each time point. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

3.3.3 mESCs are induced to differentiate unbiasedly into all three germ layers

To further understand a role of O_2 in cell specification, mESCs were used as a more physiological model of differentiation (compared to P19 cells). mESCs could be maintained undifferentiated, indefinitely, under feeder-free culture in the presence of 2i conditions and LIF, as confirmed from visualisation of the rounded and smooth-edge morphology, characteristic of pluripotent stem cells (Figure 3:9A). 2i refers to the use of 2 kinase inhibitors PD0325901 and CHIR99021, targeting MAPK and glycogen synthase kinase-3 (GSK3) respectively, which inhibit differentiation signals allowing mESCs to self-renew and induce a ground-state of pluripotency with homogeneity[349]. mESCs were unbiasedly induced to differentiate into EBs and subsequently harvested at days 3,7 and 11 for expression analysis of lineage markers by QPCR (Figure 3:9B).

Pluripotent markers Oct4 and Nanog demonstrated a time-dependent decrease in expression, indicative of cells undergoing cellular differentiation. Endodermal markers, Sox7 and Foxa2 showed a time-dependent increase in expression, reaching ~10 and ~300-fold induction respectively at day 11. Mesodermal markers, Brachyury and Gata4 were increased by day 3 and reached maximal expression by day 11, while ectodermal markers, β-III tubulin and Nestin were not increased until later time points (at day 7 and 11 respectively). These data demonstrated that the differentiation protocol could yield cells from all 3 germ layers, thereby demonstrating this cellular model of development is suitable to investigate O_2-dependent signalling in early cellular differentiation.

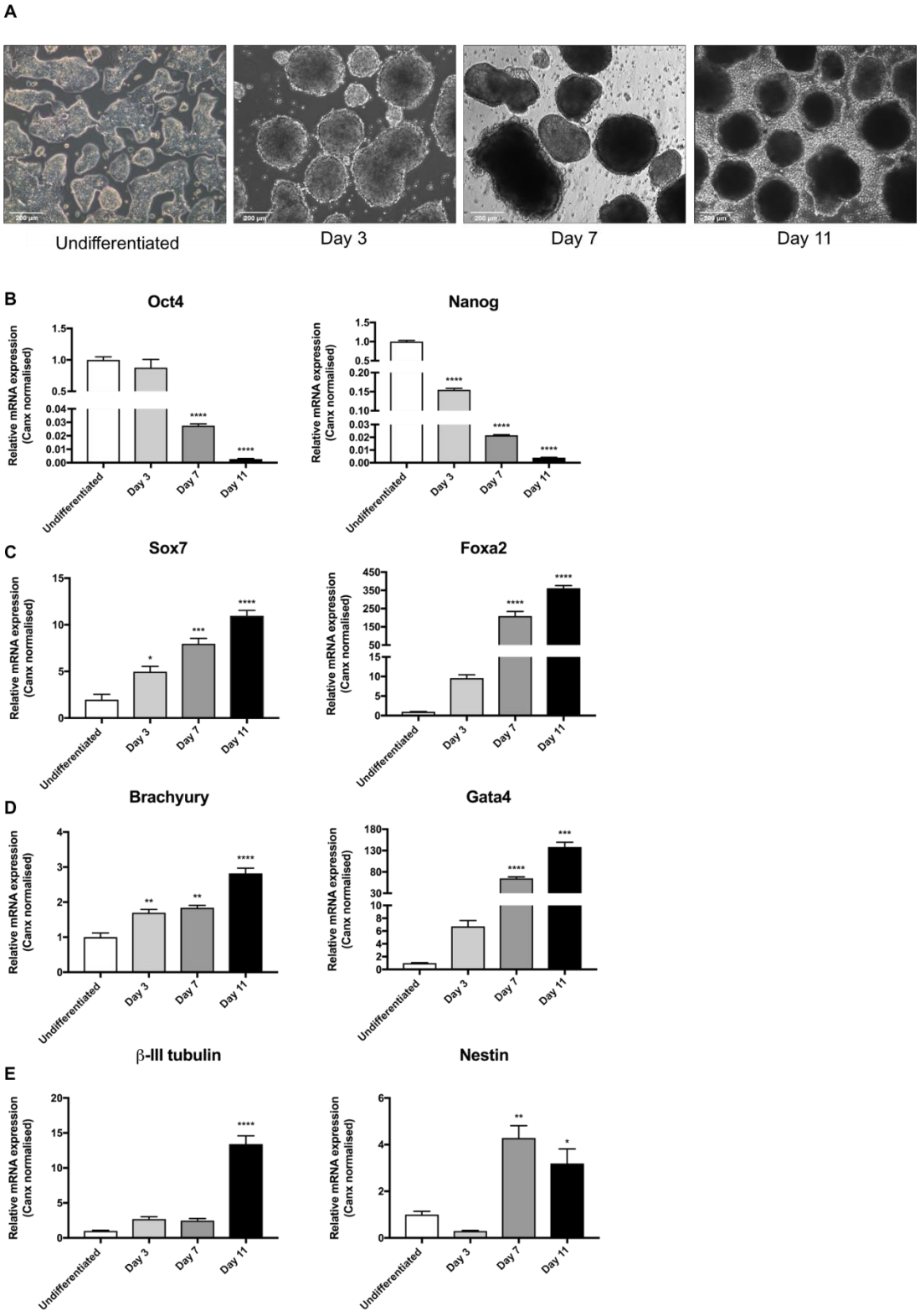


Figure 3:9 mESCs differentiate into embryoid bodies, expressing markers of all three germ layers. (A) Representative images of mESCs taken from undifferentiated to day 11 of differentiation. Expression of (B) pluripotency, (C) endodermal, (D) mesodermal and (E) ectodermal markers over an 11-day time course of differentiation. Expression data are displayed as the mean ± SEM relative to undifferentiated cells. Data are analysed by a one-way ANOVA with Dunnett's post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

3.3.4 Differentiation of mESCs under 1% O_2 skews cellular differentiation

To determine the functional effects of [O_2] upon mESC differentiation, an O_2-controlled chamber capable of exposing cells to a desired set O_2 tension (accurate range 0.5-21% O_2) was utilised (Figure 3:10A). This system was validated by the ability to detect HIF-1α protein after 4 and 24 hrs of 3% O_2 exposure in undifferentiated mESCs (Figure 3:10B). Compared to atmospheric O_2, protein expression was detected at 4 and to a lesser extent at 24 hrs, in agreement with HIF-1α showing a transient response to hypoxia[134,350,351]. Cobalt chloride ($CoCl_2$), an iron-chelating agent, which functions to inactivate the hydroxylation activity of prolyl hydroxylases, was used as a positive inducer of HIF-1α protein[352].

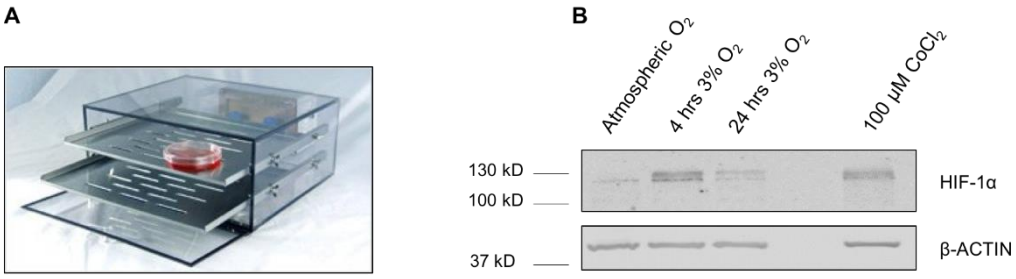


Figure 3:10 Culture of mESCs under low O_2 tension induces HIF-1α protein expression. (A) Low O_2 experiments were conducted in a BioSpherix ProOx C21 regulated C-chamber. (B) HIF-1α protein expression after 3 and 24 hrs of 3% O_2 exposure in mESCs, complete with $CoCl_2$ postive control treatment.

The effect of 1% O_2, compared to atmospheric O_2, upon mESC specification after 7 days of differentiation, was analysed by a QPCR array against a plethora of lineage markers. The expression of these markers is depicted as a fold change relative to respective atmospheric controls (Figure 3:11). Data was normalised to β-actin as the most stably expressed house-keeping gene available within the array format. Pluripotency markers were all consistently elevated in 1% O_2 above atmospheric O_2 control, suggestive of a maintained undifferentiated state and therefore delayed onset of differentiation. Endodermal markers remained largely unaffected, with no changes greater than 2-fold. A subset of mesodermal markers; Gata2[353], Brachyury and Hand1 demonstrated increased expression of approximately 7,8,18-fold respectively. However, a large proportion of other mesodermal genes failed to demonstrate greater than 2-fold changes. Ectodermal markers, and specifically the neuronal progenitor subtypes, all

demonstrated decreased expression compared to atmospheric controls. Dcx[354], a gene fundamental to neuronal migration during embryonic development, showed the largest expression change with a greater than 250-fold decrease.

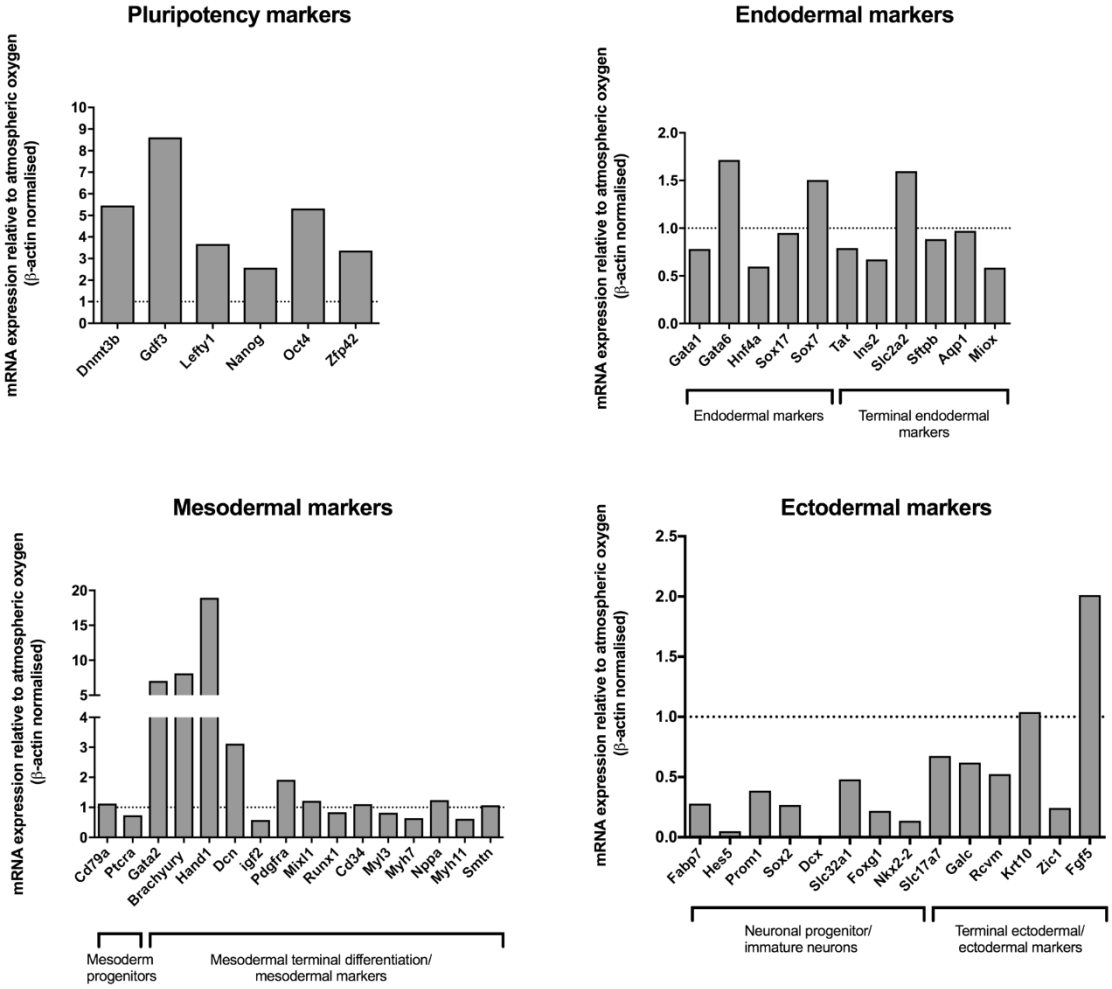


Figure 3:11 1% O_2 skews mESC differentiation. A QPCR array from pooled triplicate day 7-differentiated mESC samples exposed to 1% O_2 showing pluripotency , endodermal, mesodermal and ectodermal marker expression. Data are expressed as a fold change relative to day 7-differentiated mESCs exposed to atmospheric O_2, depicted by the dotted line at 1.

The expression changes attained from this QPCR screen were validated for a selection of genes representative of each lineage, to ensure biological significance. Additional markers were included from mesoderm and ectodermal lineages to further verify expression trends. Pluripotency markers were confirmed to be increased after 7 days of differentiation in 1% O_2, compared to atmospheric O_2 (Figure 3:12). The relative expression levels detected here by QPCR mirrored and validated the QPCR array, such that Gdf3[355] and Oct4 demonstrated the largest increase and Nanog the smallest induction. Endodermal marker expression was also consistent, displaying a significant upregulation and downregulation of Sox7 and Hnf4a[52] respectively (Figure 3:13). The increase in mesodermal expression changes for Gata2 and Hand1 were verified to be of similar fold-induction to those observed in the QPCR array, and Nppa[356] and Myh7[357] also displayed equivalent trends to those identified in the screen (Figure 3:14). However, some additional tested markers, Mesp1 and Tbx5, were significantly decreased, thereby highlighting the complexity of the effects of low [O_2] in skewing cellular differentiation. Comparatively, ectodermal markers were verified to be consistently downregulated in 1% O_2 (Figure 3:15). Supplementary markers Nestin, NeuroD[358] and β-III tubulin were also diminished in 1% O_2, compared to atmospheric O_2 controls.

These data thus validate the QPCR array findings and again demonstrate the ability of [O_2] to skew cellular differentiation patterns, which in combination with effects of DMOG in P19 cell differentiation (see section 3.3.2) highlights the potential of O_2 to act as a morphogen.

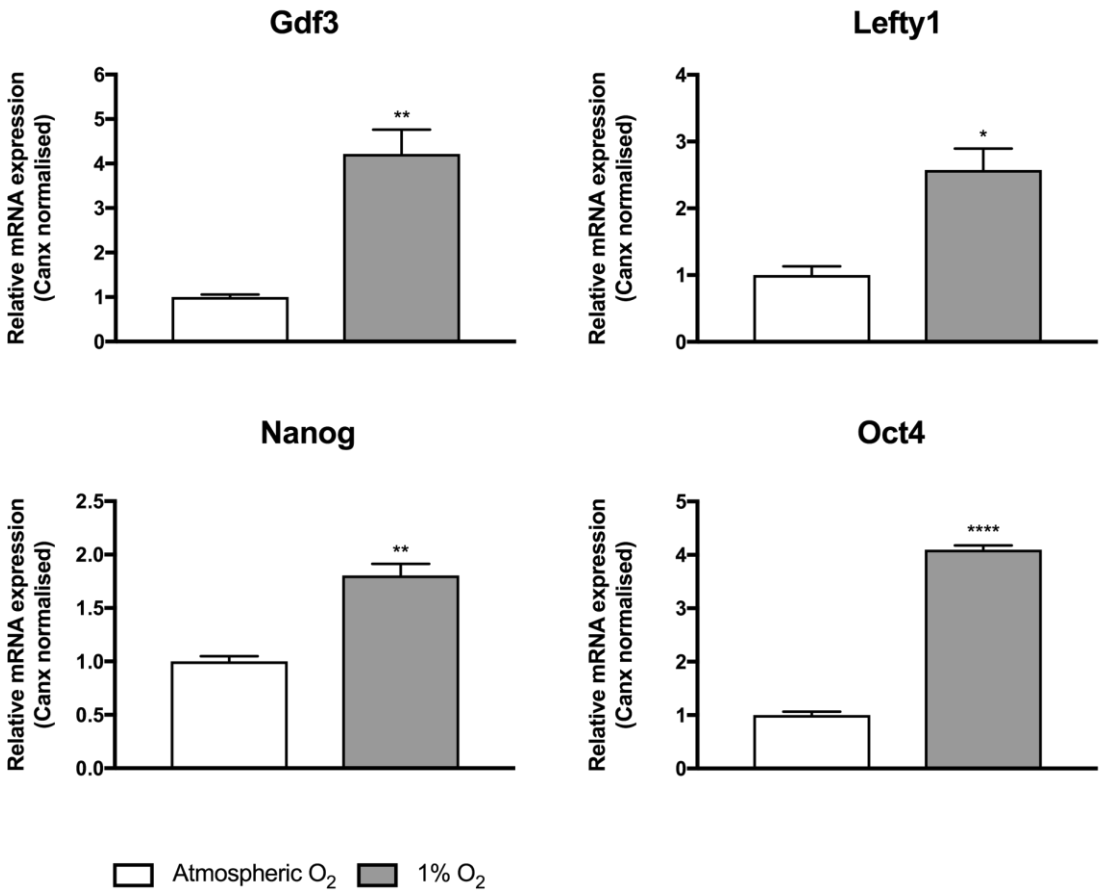


Figure 3:12 Pluripotency markers are elevated in 1% O_2 exposed day 7-differentiated mESCs. Validation of Gdf3, Lefty1, Nanog and Oct4 expression after 1% O_2 exposure, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$

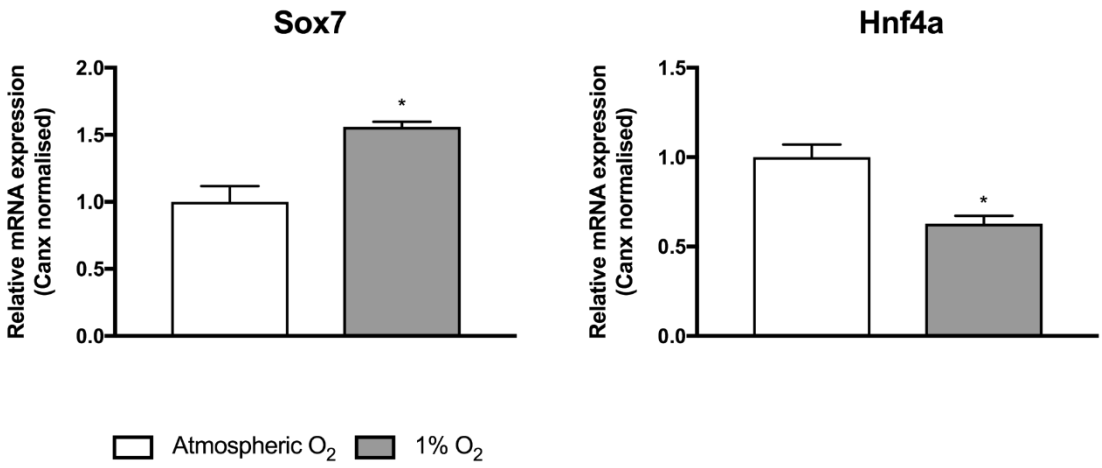


Figure 3:13 Endodermal markers show less than a 2-fold expression change in 1% O_2 exposed day 7-differentiated mESCs. Validation of Sox7 and Hnf4a expression after 1% O_2 exposure, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * $p<0.05$

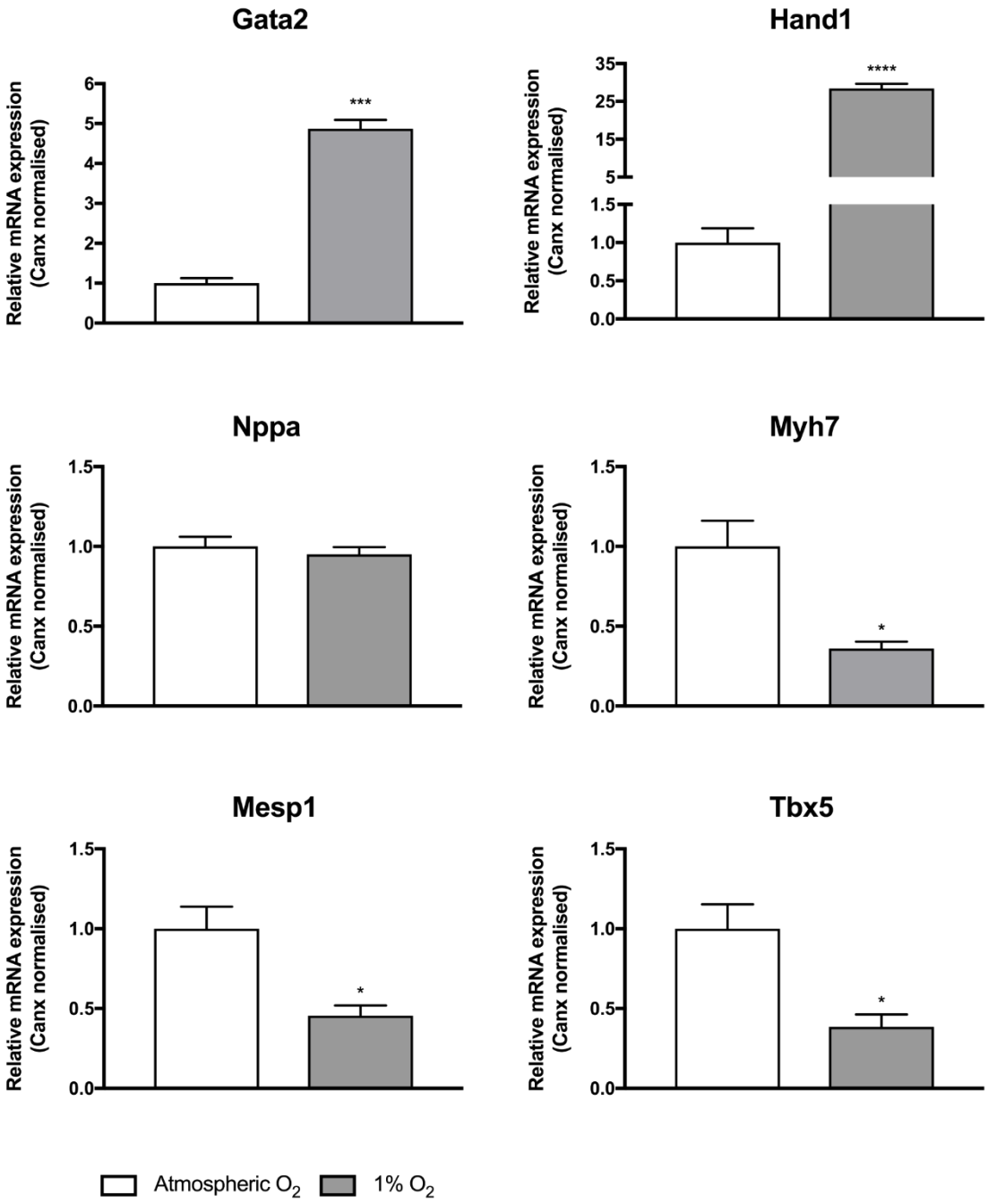


Figure 3:14 Early mesodermal markers are elevated in 1% O_2 exposed day 7-differentiated mESCs. Validation of Gata2, Hand1, Nppa and Myh7 expression, plus additional markers Mesp1 and Tbx5 after 1% O_2 exposure, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * $p<0.05$

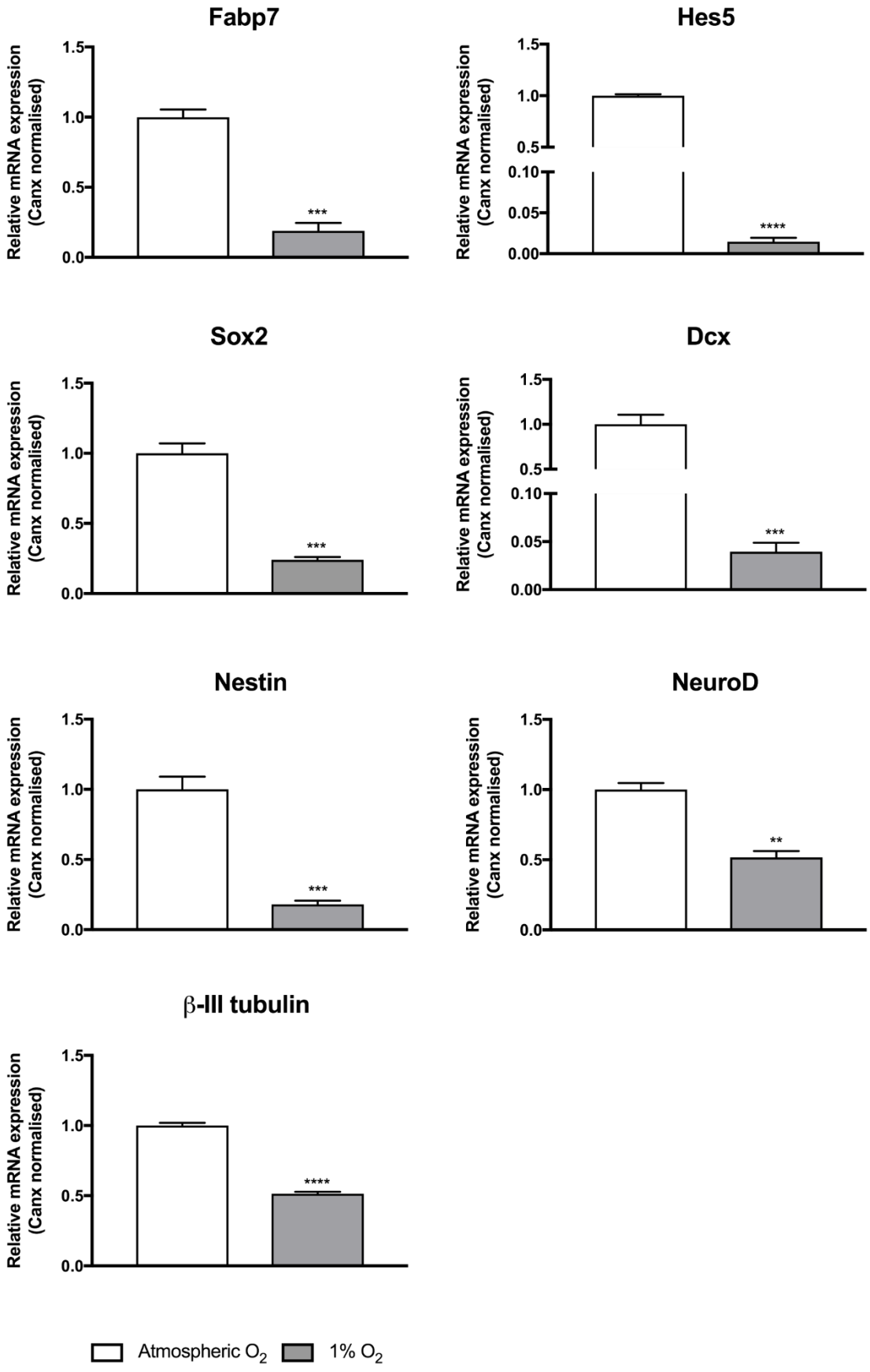

Fabp7
Hes5
Sox2
Dcx
Nestin
NeuroD
β-III tubulin
Relative mRNA expression (Canx normalised)
1.5
1.0
0.5
0.10
0.05
0.00
0.0
***
****
**
Atmospheric O_2
1% O_2

Figure 3:15 Ectodermal markers are decreased in 1% O_2 exposed day 7-differentiated mESC. Validation of Fabp7, Hes5, Sox2, Dcx expression, plus additional markers Nestin, NeuroD and β-III tubulin, after 1% O_2 exposure, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. ** p<0.01, *** p<0.001, **** p<0.0001

3.3.5 The skewing effects of 1% O_2 upon cellular differentiation are largely sustained throughout the tested time course

To investigate whether the effects of 1% O_2 upon cellular differentiation are evident at later time points, expression of a representative example of pluripotency, mesoderm and ectodermal markers were determined after 11 days of mESC differentiation in 1% O_2. The increase in pluripotency marker expression seen at day 7 was lost for Lefty1[359], but was maintained for Gdf3 and Oct4 at day 11 (Figure 3:16), albeit with a lower fold-induction. The effect of 1% O_2 upon the mesodermal markers, Hand1 and Tbx5, is consistent with that at day 7, however Mesp1 expression is no longer decreased compared to atmospheric O_2 levels (Figure 3:17). Expression of ectodermal markers remain decreased at day 11, relative to atmospheric O_2 controls (Figure 3:18), and thus are consistent with day 7.

Taken together these data suggest the effects of 1% O_2 upon cellular differentiation are largely sustained between days 7 and 11, thereby emphasising the fundamental impact of [O_2] upon cellular specification during early differentiation.

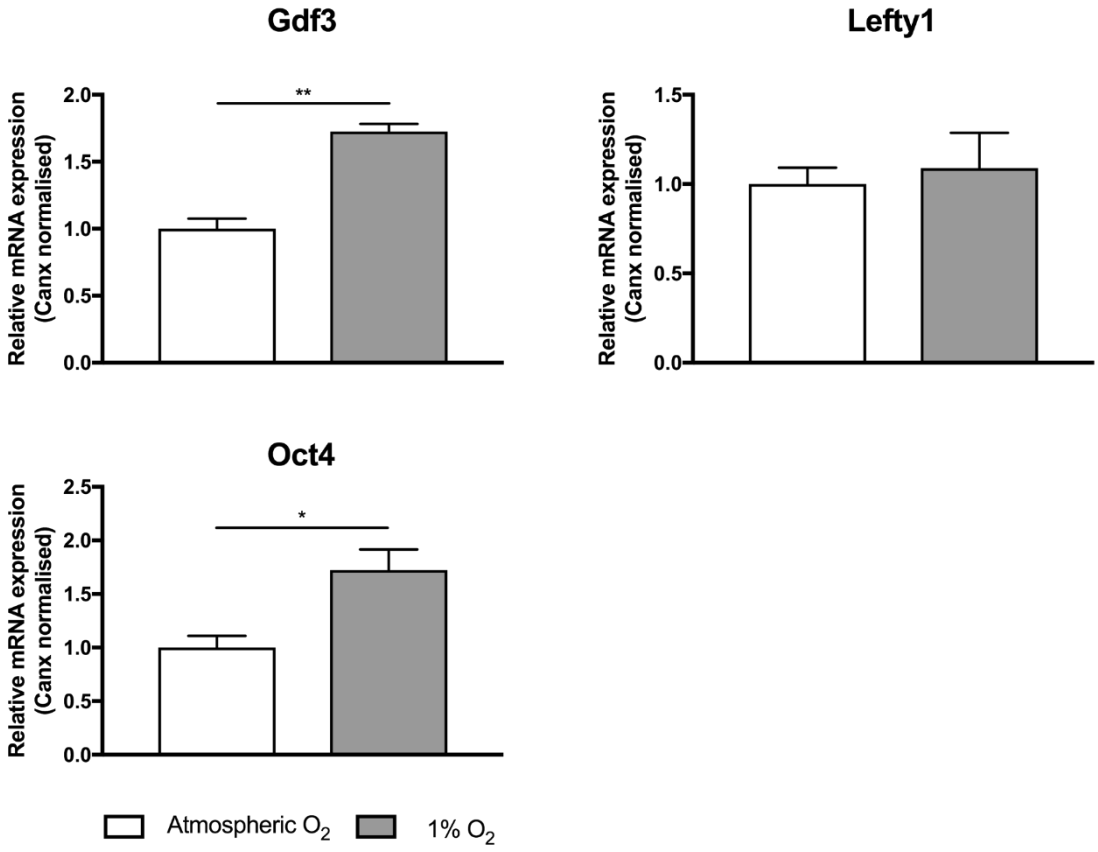


Figure 3:16 The effects of 1% O_2 upon pluripotent marker expression at day 7 is largely sustained after 11 days of mESC differentiation. Expression of Gdf3, Lefty1 and Oct4 after 11 days of mESC differentiation in 1% O_2, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * $p<0.05$, ** $p<0.01$

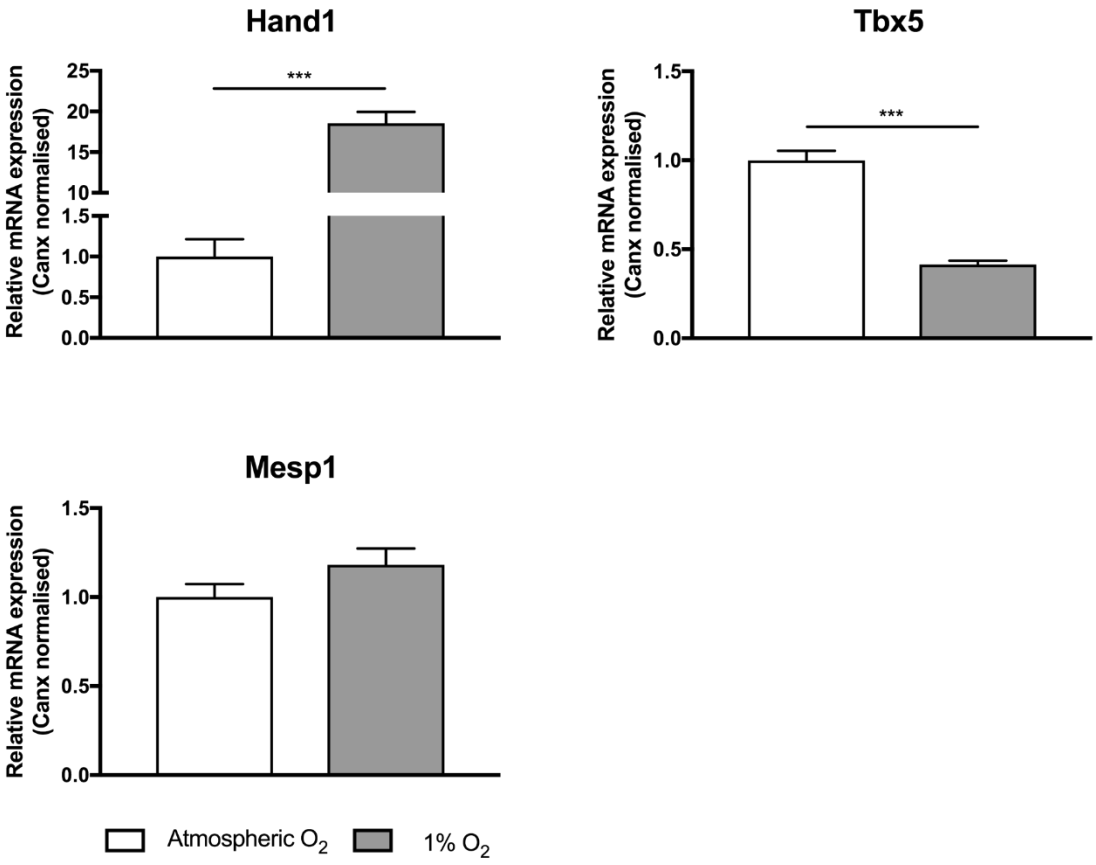


Figure 3:17 The effects of 1% O_2 upon mesodermal marker expression at day 7 is largely sustained after 11 days of mESC differentiation. Expression of Hand1, Tbx5 and Mesp1 after 11 days of mESC differentiation in 1% O_2, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. *** $p<0.001$

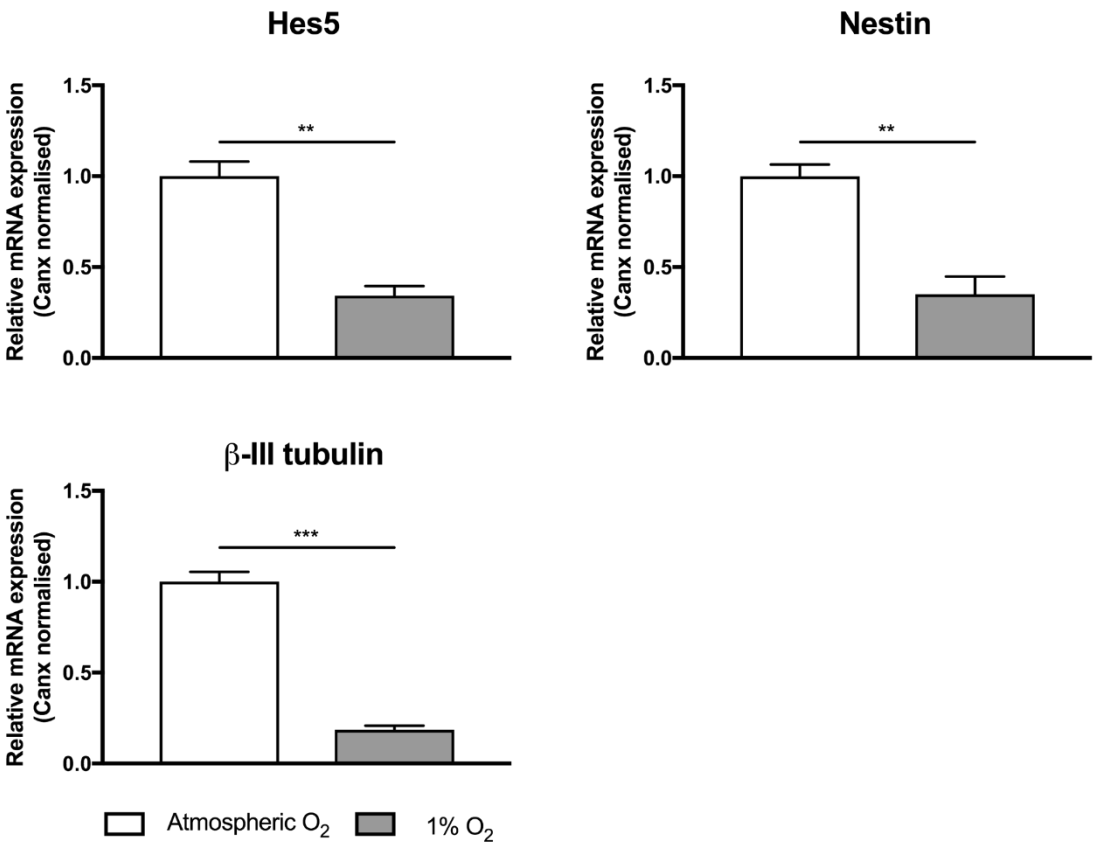


Figure 3:18 The effects of 1% O_2 upon ectodermal marker expression at day 7 is consistent after 11 days of mESC differentiation. Expression of Hes5, Nestin and β-III tubulin after 11 days of mESC differentiation in 1% O_2, relative to atmospheric O_2. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. ** $p<0.01$ *** $p<0.001$

3.3.6 Tet enzymes are differentially regulated by O_2 concentration

The requirement of O_2 as a co-factor for the enzymatic function of Tet enzymes might suggest their potential as O_2 sensors within the developing embryo. To investigate these enzymes as having a biological significance it was firstly necessary to confirm their relative expression in undifferentiated mESCs (Figure 3:19). At the mRNA level, all Tet paralogues were detectable, Tet1 was the most abundant paralogue, displaying a 9-fold and 250-fold greater expression than Tet2 and Tet3 respectively.

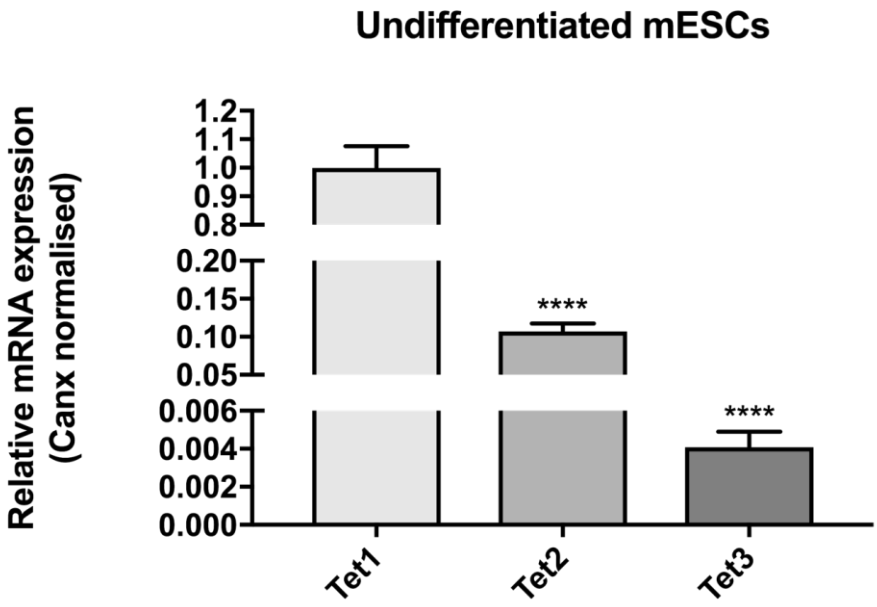


Figure 3:19 Tet1 is the most abundent Tet in undifferentiated mESCs. Relative expression of Tet1, Tet2 and Tet3 in undifferentiated mESCs. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. **** $p<0.0001$

To investigate whether the change in gene expression patterns, induced by low [O_2], during mESC differentiation could in part be mediated by Tet enzymes, it was firstly necessary to establish whether their activities are sensitive to [O_2] over physiological a range.

Plasmid constructs expressing Tet1, Tet2 and Tet3 were attained from Addgene (see Chapter 2: Material and Methods). To optimise transfection efficiency for Tet expression plasmids in mESCs, initial transfection attempts were conducted using a GFP expressing plasmid with a variety of transfection reagents (Figure 3:20) and the Amaxa™ 4D-Nucleofector™ (Figure 3:21). Induction of GFP expression 24 hrs post-transfection appeared most successful with the ViaFect™ reagent and with the CA-137 nucleofection program. However, cell viability was higher in ViaFect™-transfected cells compared to the nucleofection attempt. Both transfection methods were employed to express the Tet plasmids into mESCs. However, as determined by QPCR and Western blot analysis (data not shown), both methods failed to induce any detectable expression of the constructs, potentially attributed to the considerably larger size of the Tet plasmids, compared to the GFP plasmid, and/or the possibility the mESCs are unable to overexpress these specific proteins.

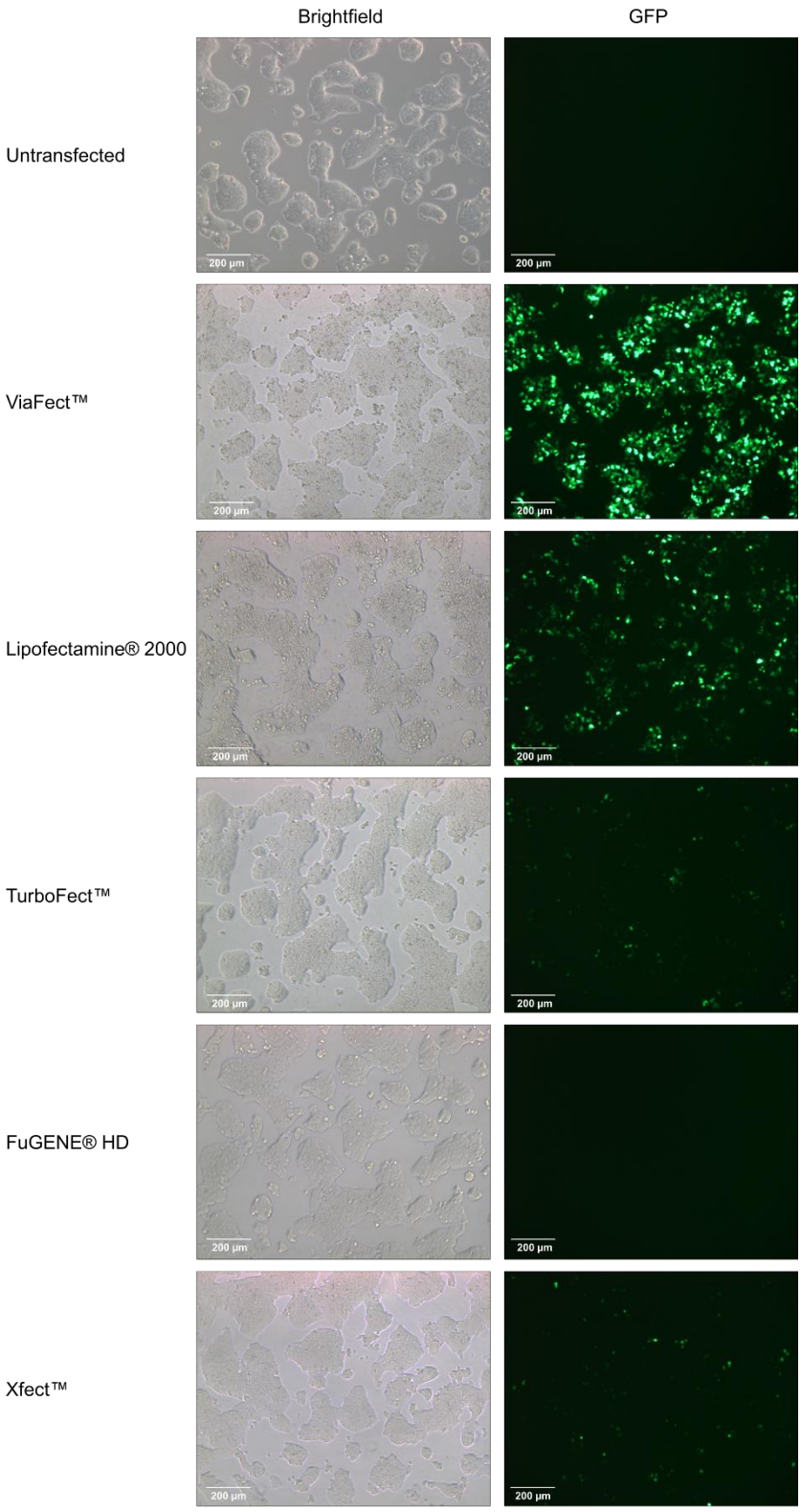


Figure 3:20 Reagent optimisation of mESCs transfections with a GFP plasmid revealed ViaFect™ demonstrated the greatest transfection efficiency. mESCs were plated onto 24-well culture plates at 100,000 cells/well 24 hrs prior to transfection of 1 µg GFP plasmid with either ViaFect™, Lipofectamine 2000®, TurboFect™, FuGENE® HD or Xfect™. Images were acquired 24 hrs post transfection.

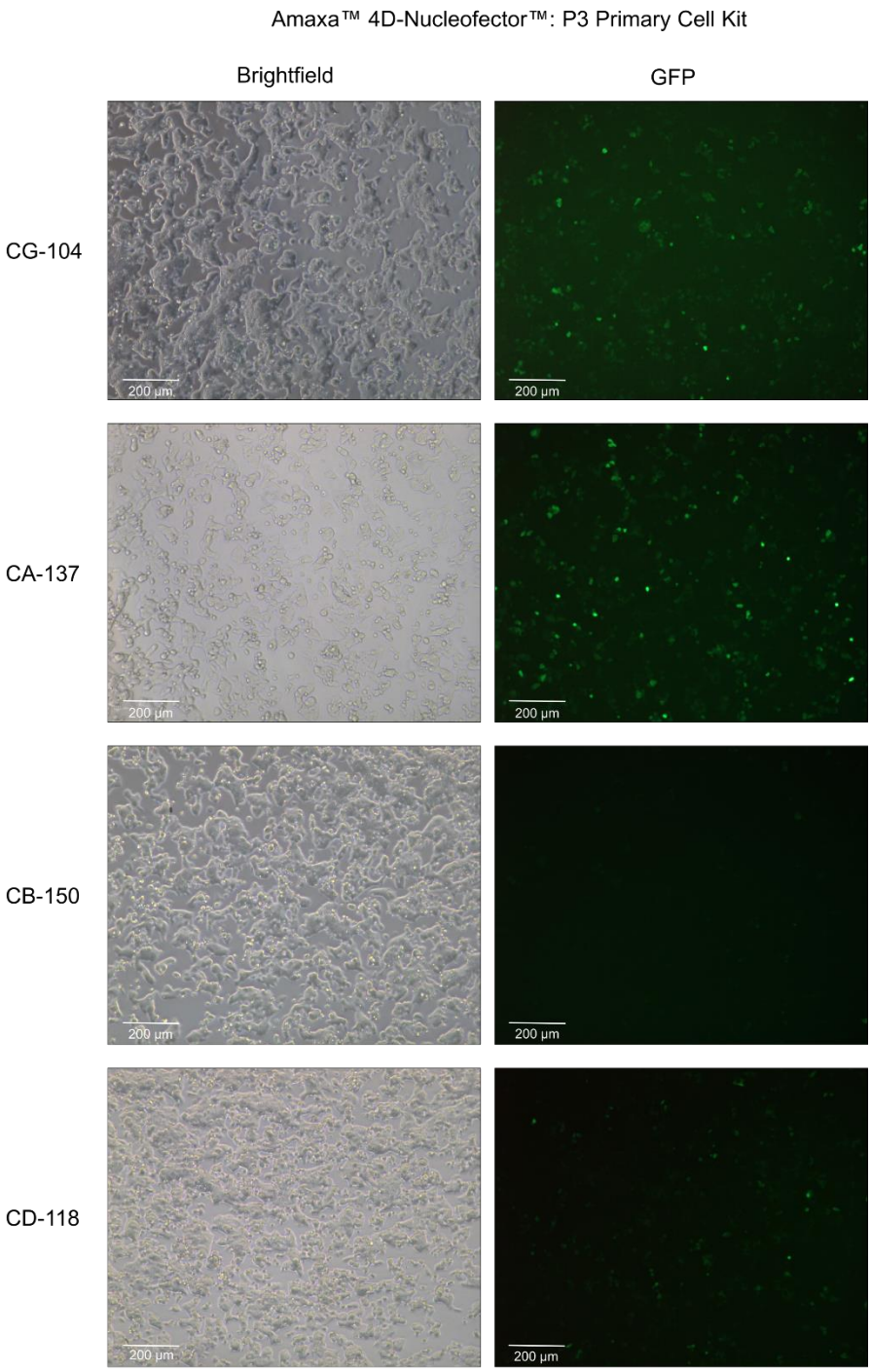


Figure 3:21 Optimisation of Nucleofection protocols for transfection of mESCs with a GFP plasmid revealed the CA-137 programme induced the highest transfection efficiency. 3x10^6 mESCs were suspended in P3 Nucleofector™ solution (Lonza) with 2 µg of pmaxGFP™ plasmid and subjected to either CG-104, CA-137, CB-150 or CD-118 programmes. Images were acquired 24 hrs post-transfection.

Due to their proven suitability for transient transfection studies, HEK cells were used as a proof-of-principle model to test whether the activities of Tet proteins might be regulated to changes in [O_2]. Tet1, Tet2 and Tet3 constructs were each successfully transiently expressed into HEK cells, as confirmed by Western blot analysis by probing against the FLAG-tag epitope of the constructs (Figure 3:22A). The result of Tet-overexpression upon 5-hmC levels was determined by dot-blot analysis. The specificity of antibodies to 5-mC and 5-hmC were firstly verified against DNA standards for unmethylated, methylated and hydroxymethylated DNA (Figure 3:22B). It was evident there was clearly no overlap in antibody selectivity, and hence these antibodies were deemed suitable for ongoing studies.

Tet-overexpressed cells were harvested for DNA 48 hrs post-transfection, and probed for 5-hmC (Figure 3:22C). Significant inductions in 5-hmC levels above those in pcDNA-transfected control cells was evident for each construct. Tet2 overexpression demonstrated the largest increase in 5-hmC, approximately 30-fold greater than control.

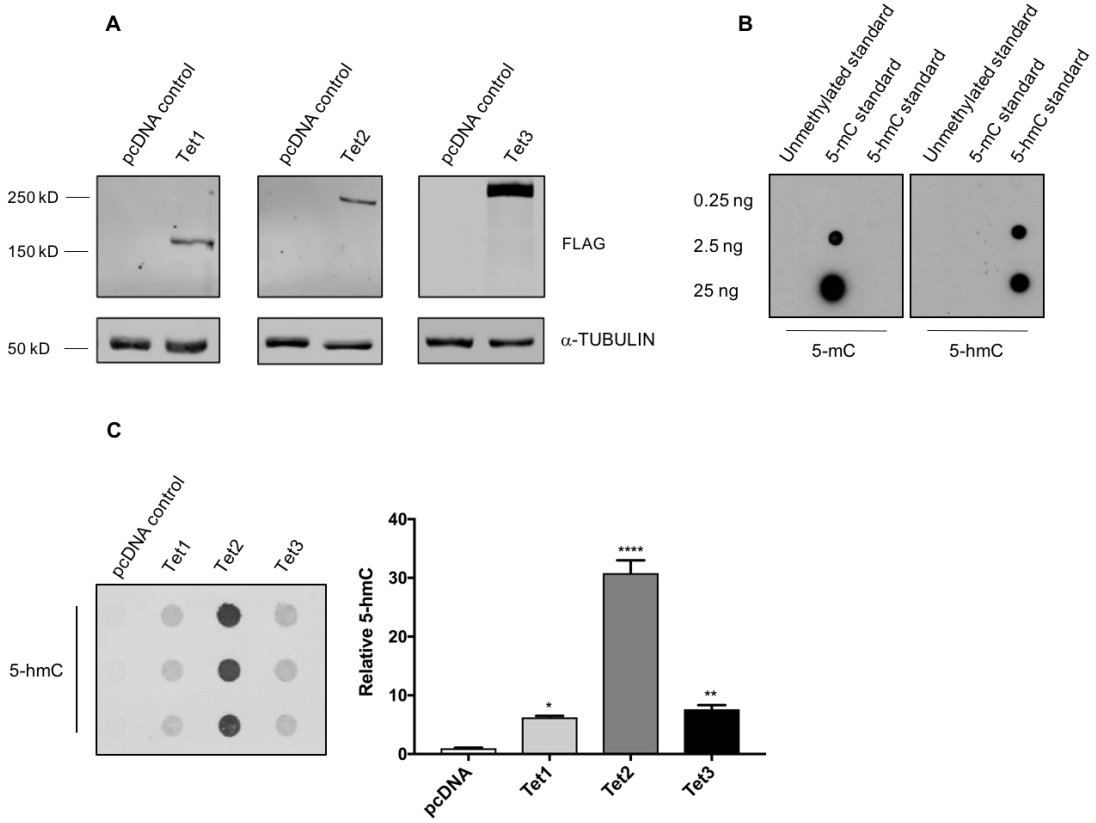
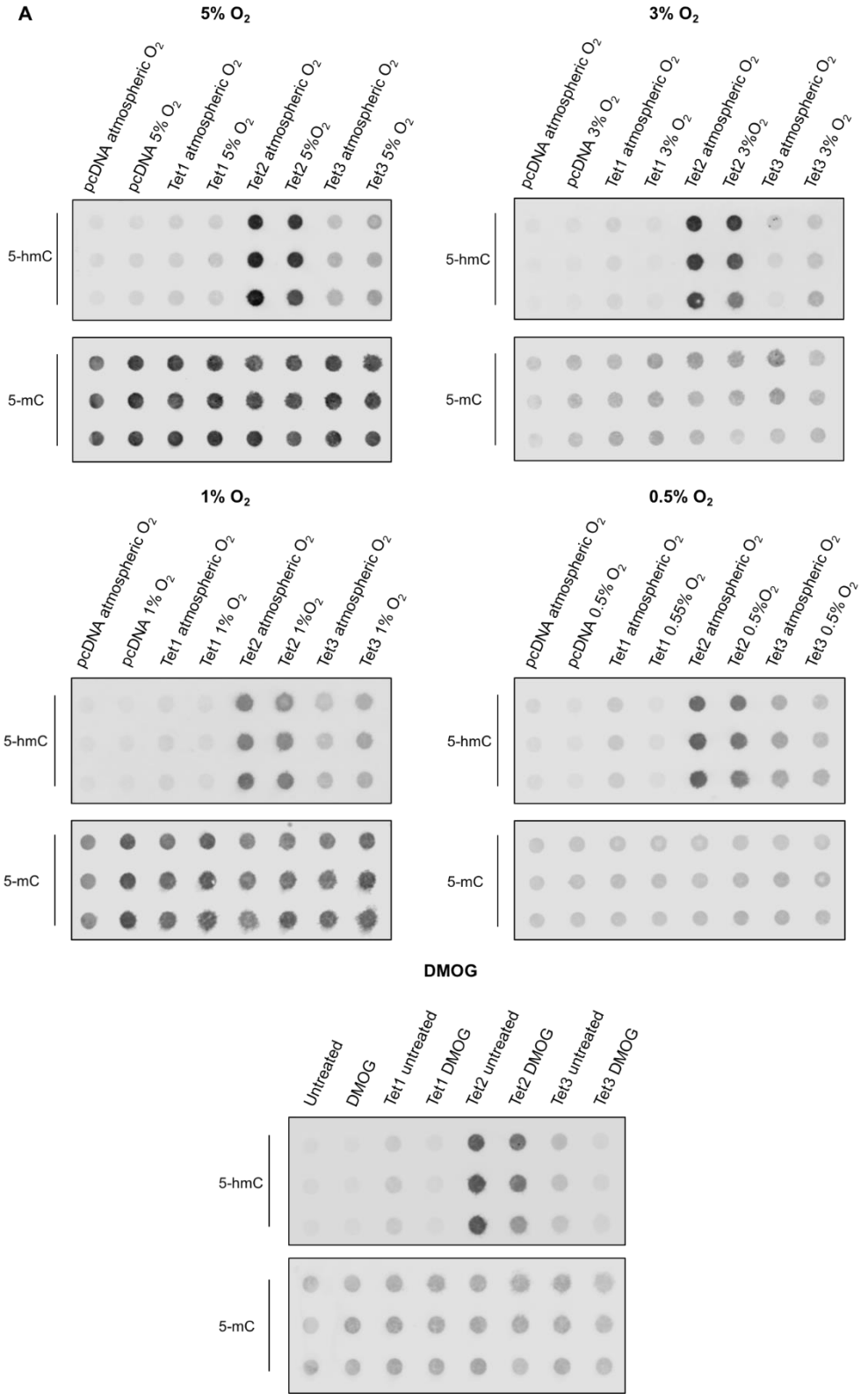


Figure 3:22 Overexpression of Tet proteins in HEK cells increased 5-hmC above pcDNA control transfected cells. (A) The ability to overexpress each Tet protein in HEK cells was assessed by Western blot analysis, probing against the FLAG-tag epitope of the constructs. (B) The specificty of antibodies against 5-mC and 5-hmC were asessed against known DNA standards (unmethylated DNA, 5-mC DNA and 5-hmC DNA) by dot-blot analysis. (C) The functionality of Tet overexpression constructs was determined by 5-hmC detection by dot-blot analysis, and levels quantified and expressed relative to pcDNA control transfected cells. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

Tet-overexpressing HEK cells were cultured in decreasing O_2 concentrations, over a range of 5, 3, 1 and 0.5% O_2. These O_2 concentrations were chosen to represent graded levels of O_2 that are physiologically relevant in the developing embryo. Levels of 5-hmC and 5-mC, detected firstly by dot-blot, were used as a measure of Tet activity. Each dot-blot (Figure 3:23A) shows 5-hmC and 5-mC levels within HEK cells overexpressing each Tet enzyme under atmospheric O_2 and the tested [O_2]. In addition, 1 mM DMOG (at atmospheric O_2) was used as a positive control for inhibition. Visual interpretation of 5-hmC expression clearly demonstrates that Tet1 activity is decreased at 3, 1, and 0.5% O_2, compared to respective atmospheric O_2 controls. Tet2 appeared to be largely unaffected at all O_2 concentrations, whereas Tet3 at 3% O_2 unexpectedly increased 5-hmC. 5-mC levels appeared largely unaffected by O_2 concentrations and were used as an internal control for DNA loading to produce a ratio of 5-hmC:5-mC for quantification (Figure 3:23B). 5-hmC:5-mC were expressed as percentage of activity compared to respective atmospheric O_2 controls, represented by the dotted line at 100%. Overall, these results demonstrate that Tet1 activity was unaffected at 5% O_2, but significantly inhibited at 3% O_2 and less, to approximately 40% of maximal activity. Tet2 activity was only significantly affected at 0.5% O_2, to around 70% of maximal activity, equivalent to the effect with DMOG treatment. Based on the requirement of Tet enzymes for O_2 as a co-substrate it would be hypothesised that activity would be decreased in low [O_2]. However, as stated above intriguingly an approximate 3-fold increase in Tet3 activity at 3% O_2 was observed. Other tested O_2 concentrations had no significant effect on Tet3 activity, and only DMOG treatment resulted in an approximate 50% inhibition.

To verify these findings, mass-spectrometry was used as a more sensitive approach for 5-hmC and 5-mC detection (Figure 3:24). 3 and 0.5% O_2-exposed cells were analysed for C, 5-mC and 5-hmC levels. Here, 5-hmC levels were normalised to total C (combined C, 5-mC and 5-hmC) and expressed as a percentage of activity, compared to corresponding atmospheric O_2, depicted by the dotted line at 100%. This analytical method decreased the variability between replicates and thus in combination with dot-blot findings confirmed that inhibition of Tet1 was the most sensitive to changes in [O_2] over this tested range, compared to Tet2 and Tet3. It should be noted that mass-spectrometry analysis did highlight discrepancies, compared to dot-blot analysis, but only in the pcDNA-transfected control samples. The increase in 5-hmC at 3% O_2 reported from the mass-spectrometry is likely down to encompassing total C normalisation within the cell, most probably reflected by increased endogenous Tet3 activity.


A
5% O2
pcDNA atmospheric O2
pcDNA 5% O2
Tet1 atmospheric O2
Tet1 5% O2
Tet2 atmospheric O2
Tet2 5%O2
Tet3 atmospheric O2
Tet3 5% O2
5-hmC
5-mC
3% O2
pcDNA atmospheric O2
pcDNA 3% O2
Tet1 atmospheric O2
Tet1 3% O2
Tet2 atmospheric O2
Tet2 3%O2
Tet3 atmospheric O2
Tet3 3% O2
5-hmC
5-mC
1% O2
pcDNA atmospheric O2
pcDNA 1% O2
Tet1 atmospheric O2
Tet1 1% O2
Tet2 atmospheric O2
Tet2 1%O2
Tet3 atmospheric O2
Tet3 1% O2
5-hmC
5-mC
0.5% O2
pcDNA atmospheric O2
pcDNA 0.5% O2
Tet1 atmospheric O2
Tet1 0.55% O2
Tet2 atmospheric O2
Tet2 0.5%O2
Tet3 atmospheric O2
Tet3 0.5% O2
5-hmC
5-mC
DMOG
Untreated
DMOG
Tet1 untreated
Tet1 DMOG
Tet2 untreated
Tet2 DMOG
Tet3 untreated
Tet3 DMOG
5-hmC
5-mC

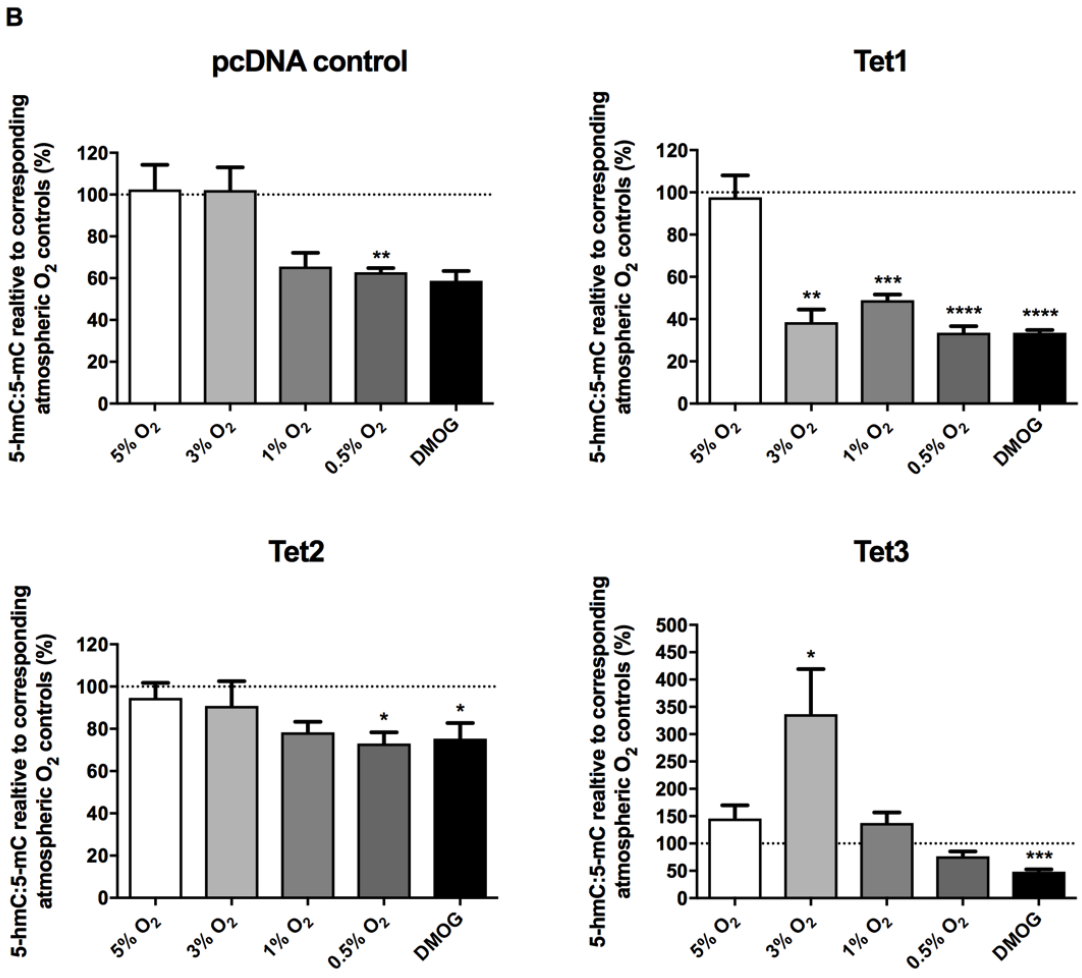


Figure 3:23 Tet proteins displayed differential regulation by O_2 concentration. (A) pcDNA control or Tet1, Tet2 and Tet3 expressing plasmids were transfected into HEK cells and subjected to decreasing O_2 concentrations (5,3,1 and 0.5% O_2) for 48 hrs, and harvested for 5-hmC and 5-mC detection by dot-blot analysis. Cells were treated with 1 mM DMOG, under atmospheric O_2, as a postive inhibitor control. (B) Quantifcation of dot blots, displayed as the ratio of 5-hmC:5-mC, expressed as a percentage of activity compared to atmospheric control (depicted by the line at 100%). Data are shown as the mean ± SEM and analysed by unpaired t-test between atmospheric O_2 and the tested condition. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

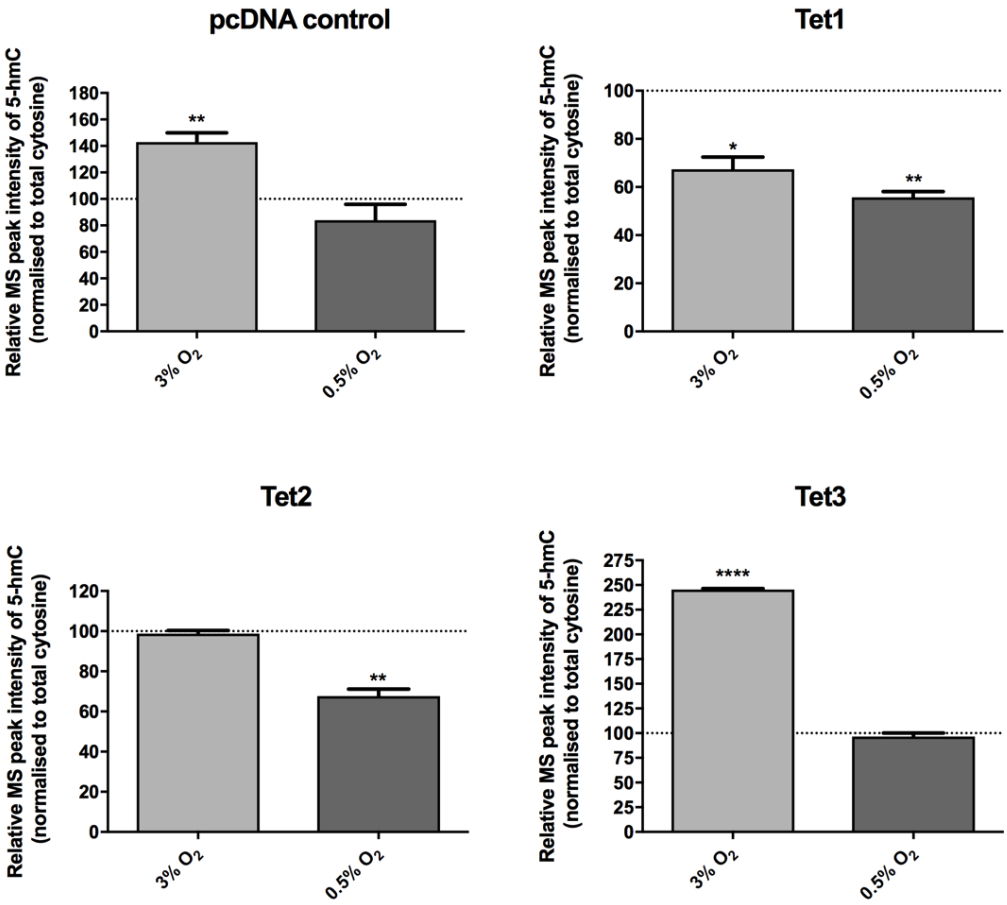


Figure 3:24 Mass-spectrometry analysis confirms the differential regulation of Tet enzymes by O_2 concentration. DNA samples from pcDNA control and Tet1, Tet2, and Tet3 expressing HEK cells, subjected to 3 and 0.5% O_2 were analysed for C, 5-mC and 5-hmC by mass spectrometry. Data is normalised to total C and expressed as a percentage of activity compared to corresponding atmospheric O_2 control (depicted by the line at 100%). Data are shown as the mean ± SEM and analysed by a unpaired t-test between atmospheric O_2 and the tested condition. * $p<0.05$, ** $p<0.01$

3.3.7 O_2 concentration regulates cellular metabolite levels

In addition to O_2, Tet enzymes also require 2-OG as a co-substrate and are known to be antagonised by the sterically similar citric acid cycle metabolites, succinate and fumarate[303]. Culture under differing O_2 concentrations will affect cellular metabolism and hence may therefore influence Tet activity. Thus, to investigate a potential cause for the unexpected increase in Tet3 activity seen in response to 3% O_2, the metabolic profile of HEK cells was determined by ^{1}H-NMR after culture in 3% and 1% O_2 (Figure 3:25). Lactate levels, indicative of glycolysis activation, were increased at 1% O_2, above atmospheric O_2. Glutamate levels, intrinsic for the formation of the 2-OG citric acid cycle intermediate, were significantly reduced in both 3% and 1% O_2. These data suggest that 2-OG levels would likely be reduced and would therefore be predicted to inhibit Tet activity. However, succinate levels were also decreased at 3% O_2, but not at 1% O_2, compared to atmospheric O_2. Hence, there might be potentially less Tet inhibition by succinate at 3% O_2 compared to 1% O_2, and this may account for the increased activity of Tet3. No significant differences were found in fumarate levels between 3% and 1% O_2.

Irrespective of the specific mechanisms underlying the biphasic response of Tet3 activity to differing O_2 concentrations, the activities of Tet enzymes are demonstrated here to be regulated by O_2 levels. Notably, Tet1 activity is specifically sensitive to changes in O_2 concentrations that are physiologically relevant within the developing embryo and thus may have potential to act as an O_2 sensor.

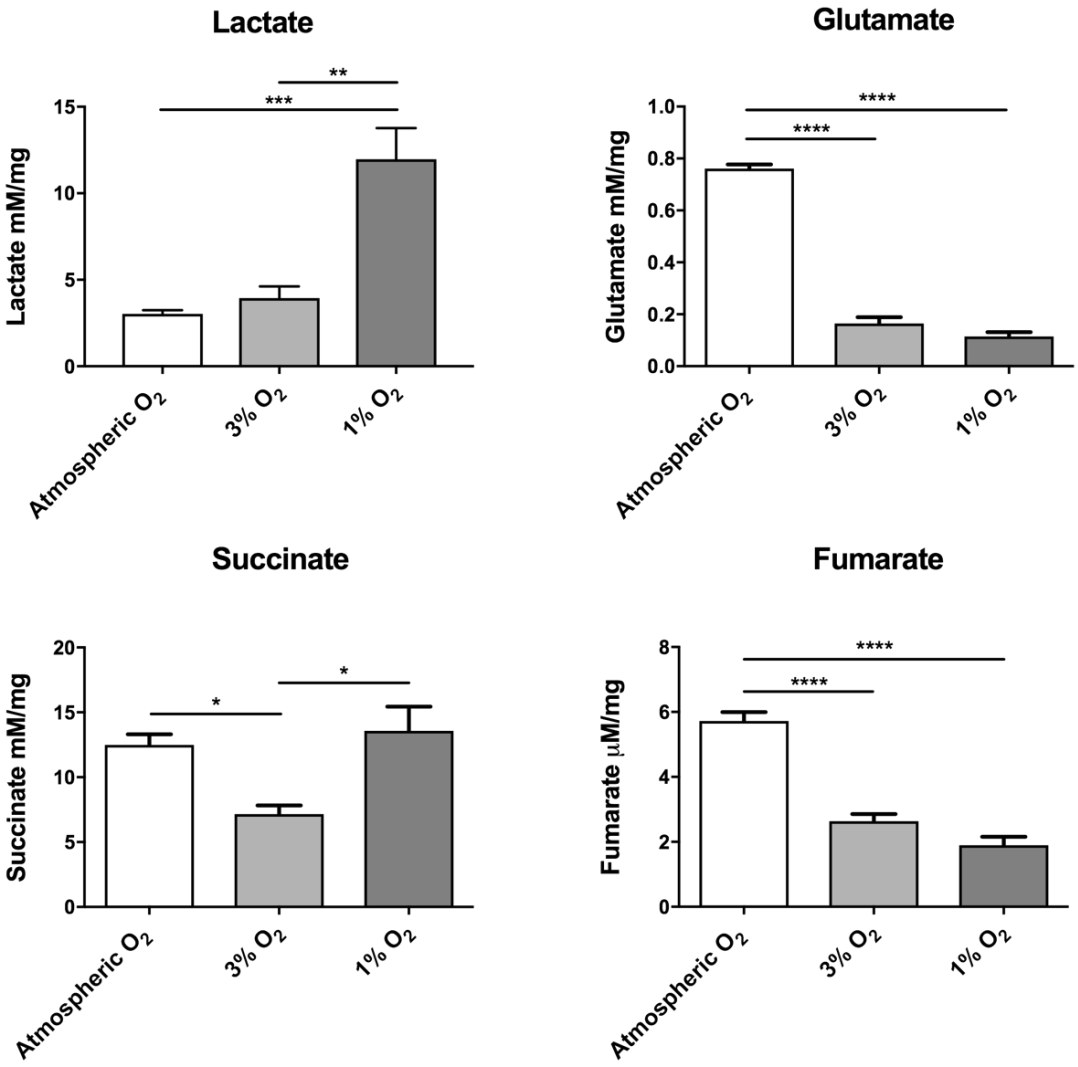


Figure 3:25 Differences in succinate levels between 3 and 1% O_2 may in part explain the elevated Tet3 activity at 3% O_2. HEK cells were subjected to atmospheric, 3% and 1% O_2 for 24 hrs and cellular metabolite levels determined by ^{1}H-NMR. Data are expressed as the mean ± SEM from 5 repeats and normalised to total protein. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

3.3.8 Generation of stable Tet-depleted mESCs

The activities of Tet enzymes appear differentially regulated by [O_2], as shown previously in section 3.3.6. However, to investigate directly whether Tets are in part responsible for mediating mRNA expression changes induced by 1% O_2, as presented in section 3.3.4, each Tet paralogue was genetically ablated. Genetic ablation of Tet1, Tet2 and Tet3 in mESCs was performed by shRNA lentiviral transduction and stable knockdown (KD) cell lines were generated following positive selection with puromycin

A dose response for antibiotic selection was firstly conducted by exposing undifferentiated mESCs to increasing puromycin concentrations to determine the minimal concentration to induce cell death after 72 hrs (Figure 3:26). Complete cell death after this time was achieved with 1 µg/ml puromycin.

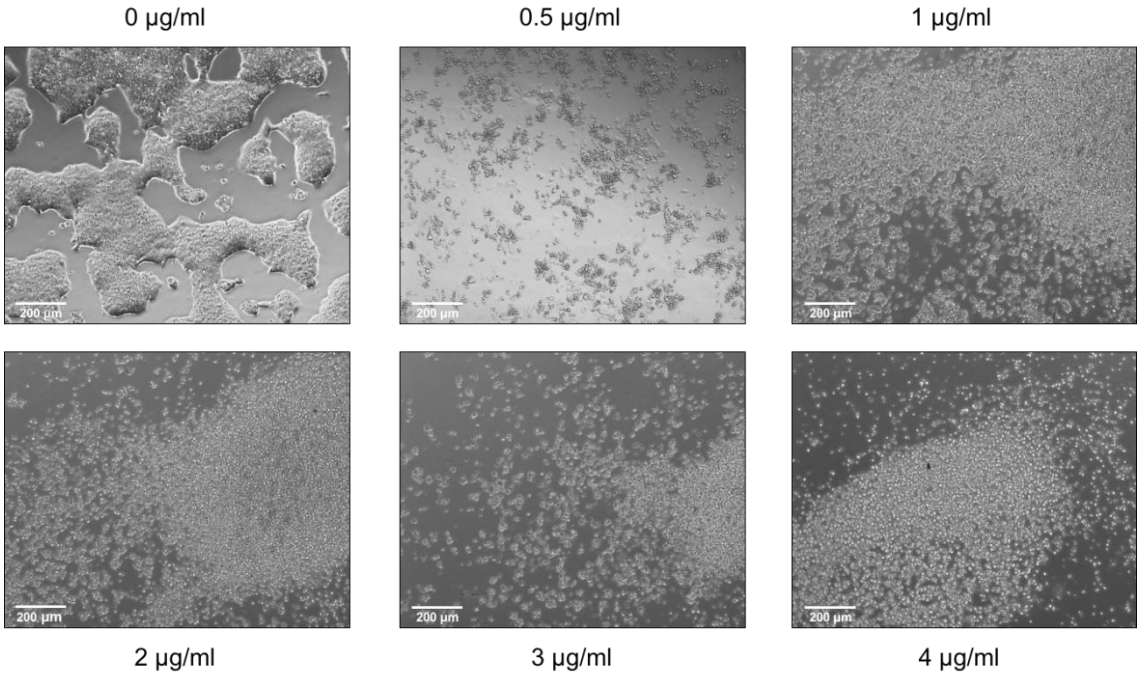


Figure 3:26 A dose response of puromycin treatment revealed 1 µg/ml was sufficient to induce mESC death after 72 hrs. Undifferentiated mESCs were treated with increasing concentrations of puromycin (0.5-4 µg/ml) for 72 hrs, 1 µg/ml was selected as the minimal concentration to induce cell death after this time.

2 independent shRNA clones (see Chapter 2: Materials and Methods) against each Tet enzyme were assessed to determine their gene silencing ability by QPCR analysis (Figure 3:27) (Due to a high level of Tet3 ablation with clone 1, the use of clone 2 was not necessary). The greatest level of silencing of Tet1 was achieved with clone 2 (~75%) and for Tet2 this was clone 1 (~55%). Tet3 demonstrated the greatest level ablation of ~95% compared to shRNA control. Clones that produced the greatest level of Tet silencing were used to produce stable cell lines for use in future studies.

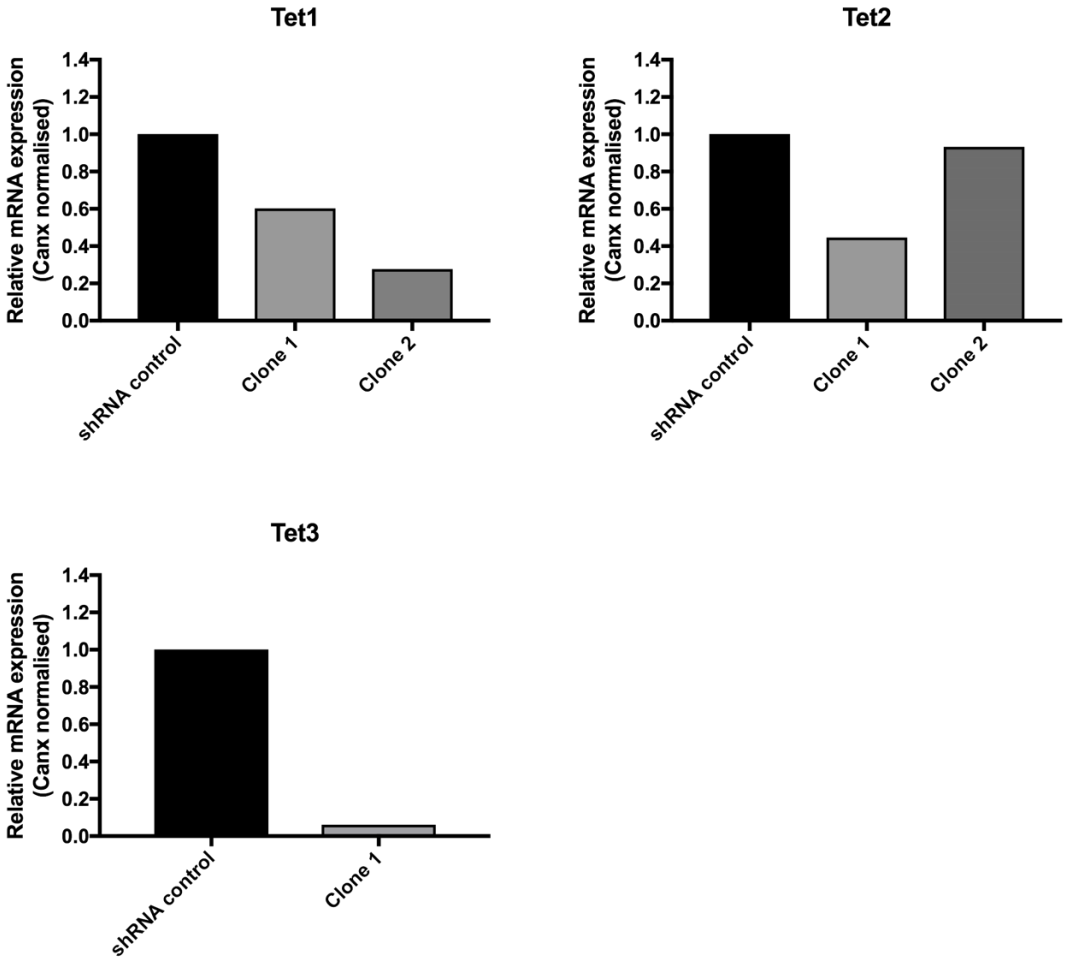


Figure 3:27 Genetic ablation of Tet1, Tet2 and Tet3 in undifferentiated mESCs. Assessment of genetic silencing capabilities of lenti-viral shRNA clones for Tet1, Tet2, Tet3 in mESCs was performed by QPCR analysis. mESCs were firstly infected with lenti-viral particles and after 72 hrs postively transduced cells were selected and expanded with 1μg/ml puromycin containg medium. Data are expressed relative to shRNA control infected cells.

To determine whether genetic silencing at the mRNA level correlates to protein expression, attempts were made to confirm ablation of Tet1, Tet2 and Tet3 by Western blot analysis (Figure 3:28). The Tet1 protein band appeared difficult to distinguish, however based on the assumption that the protein band seen by probing the Tet overexpression construct with FLAG would be of a similar size (see Figure 3:22C) (the FLAG tag is small and therefore would not alter the protein size significantly), the endogenous Tet1 was depicted by the arrow. Tet1 protein level showed a 20% decrease compared to a 75% reduction at the mRNA level. Tet2 ablation resulted in an equivalent loss at the protein level, depicted by the arrow, to that at the mRNA level. Virtually no protein was detected for Tet3, in agreement with the magnitude of ablation at the gene level. However, it should be noted that although Tet3 mRNA is the lowest expressed Tet paralogue in undifferentiated mESC it is readily detectable at the protein level. This may be due to several possibilities including the mRNA may be degraded after translation, translation itself may be upregulated and the protein may have a long half-life due to a high stability allowing it to become accumulated, whereas the mRNA has a high turnover.

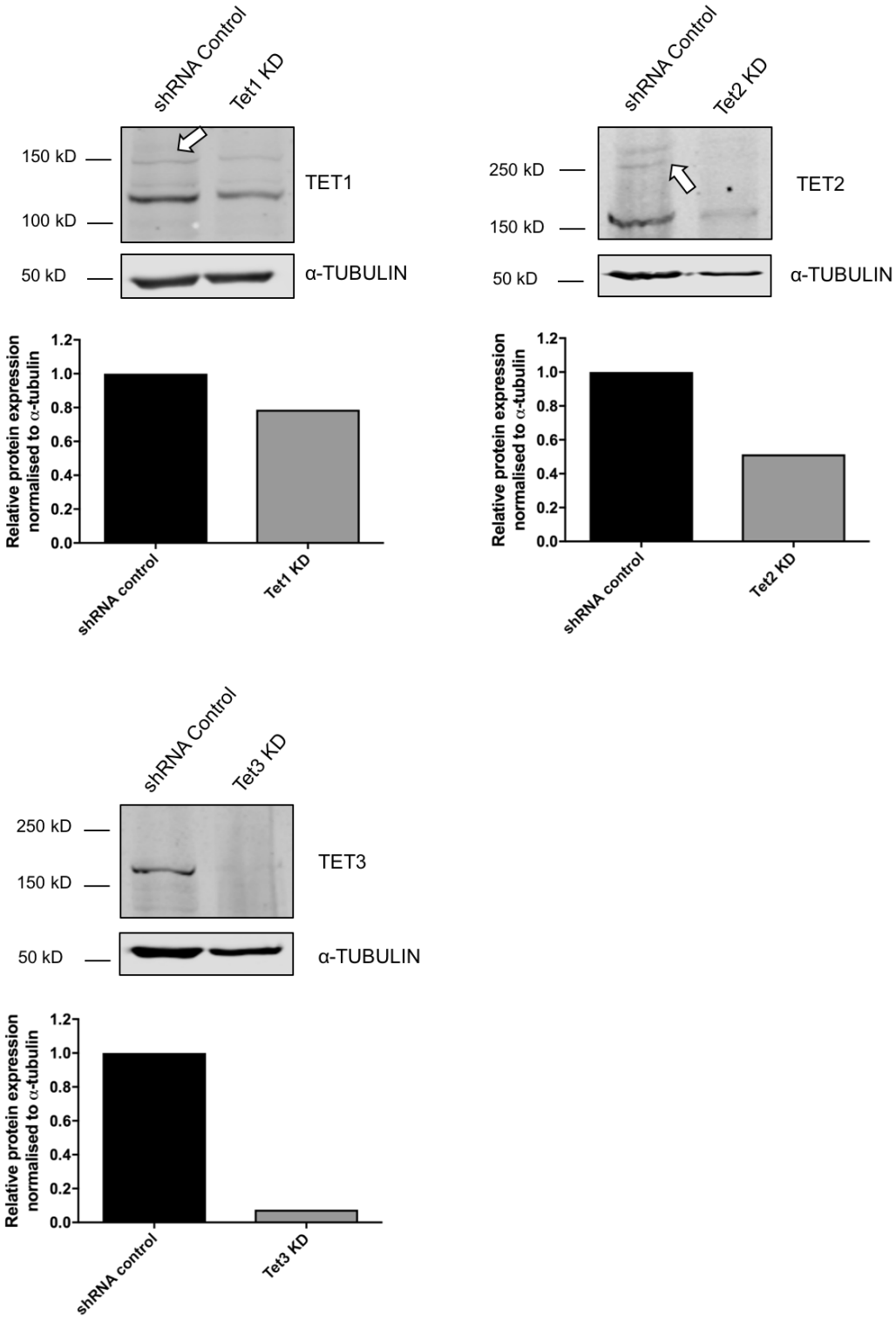


Figure 3:28 Genetic silencing of Tets at the mRNA level correlates with the loss of protein expression. Positively selected Tet1, Tet2 and Tet3 lenti-viral shRNA transduced mESCs were harvested for protein analysis by Western blot. Where necessary the correct protein bands are marked with a white arrow. Quantification of protein expression is expressed relative to shRNA control. KD: knockdown.

3.3.9 Genetic ablation of each Tet enzyme has potential compensatory effects on other paralogues

To distinguish whether silencing of each Tet enzyme affects expression of the other paralogues, mRNA levels were assessed in undifferentiated mESCs (Figure 3:29). Successful ablation of each Tet paralogue was again confirmed. However, it was apparent that Tet2 expression was elevated in Tet3 KD and Tet3 expression was decreased in Tet2 KD. This suggests a potential inter-relationship between the Tet enzymes, highlighting potential compensatory effects that should be considered when interpreting Tet silencing studies.

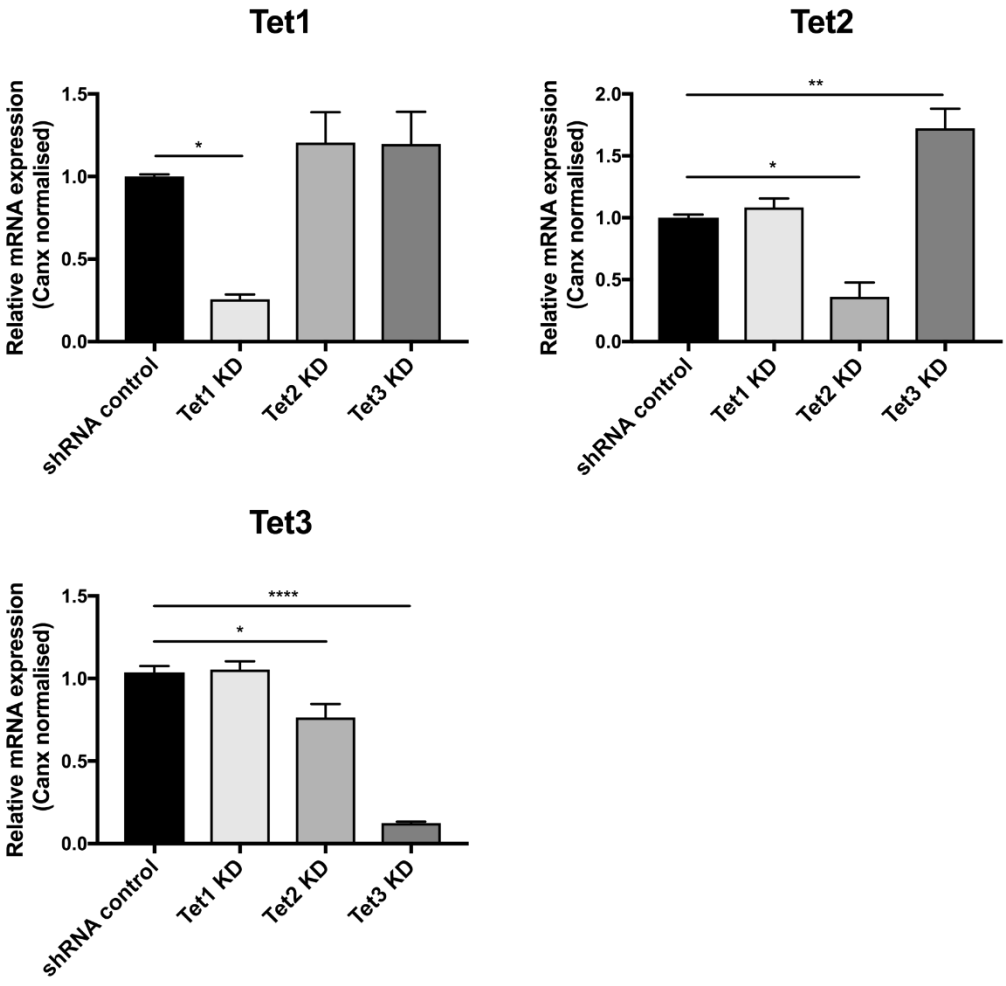


Figure 3:29 Compensatory effects of Tet ablation on other Tet paralogues in undifferentiated mESCs. Expression of Tet1, Tet2 and Tet3, relative to shRNA control was assessed in Tet1, Tet2 and Tet3 genetically abalted mESCs by QPCR analysis in undiffereniated cells. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Tukey's post-hoc analysis. *p<0.05, **p<0.01, **** p<0.0001. KD: knockdown.

3.3.10 Tet knockdown inhibits mESC differentiation

The ability of Tet enzymes to in part be responsible for the changes in lineage marker expression reported for mESC differentiation under 1% O_2 (thereby identifying differential roles in cellular specification) was investigated. Tet1-, Tet2- or Tet3-ablated mESCs were induced to differentiate for 7 days and a QPCR array against a range of lineage markers, as performed in Figure 3:11, was conducted (Figure 3:30). Data is expressed as a fold change relative to respective shRNA controls and normalised to the β-actin house-keeping gene, selected as the most stably-expressed gene within the array format. All pluripotency markers were elevated in Tet1-, Tet2-, Tet3-ablated cells, with Tet3 silencing inducing the most pronounced effect. These effects were validated, which confirmed Tet1 ablation trended to increase expression of all pluripotent markers, but only achieved significant induction of Dnmt3b[360] and Oct4 (Figure 3:31). Tet2 silencing resulted in a significant increase in all tested markers, except Dnmt3b (Figure 3:32). Tet3 ablation, as first seen from the QPCR array, maintained the largest significant fold induction of all tested pluripotency genes, compared to Tet1 and Tet2 KD mESCs (Figure 3:33). Hence these effects are consistent with the increase in pluripotency marker expression established when differentiating cells were cultured in 1% O_2 (Figure 3:11).

Endodermal markers showed little change in expression with either Tet1-, 2- or 3-ablation, and thus was similar to effects induced by 1% O_2 (Figure 3:11). The majority of endodermal markers showed the same trend with silencing of either Tet paralogue. However, in the case of Foxa1[361], Gata1[362] and Hnf4a there appeared some minor, yet differential effects between ablation of Tet1, 2, 3. Expression of mesodermal markers were largely unchanged (less than a 2-fold change), compared to shRNA control. Ablation of either Tet enzyme failed to mimic the large induction of the early mesodermal markers Gata2, Brachyury, Hand1 and Dcn[363] seen with 1% O_2 (Figure 3:11). Instead, Hand1 expression was in fact substantially and consistently reduced with all Tet silenced paralogues. Overall, largely similar effects on expression of each mesodermal marker was seen in mESCs ablated of either Tet enzyme. Although some clear differences were recorded, such as the increased Gata2 expression in Tet1-silenced cells and reduced expression of Ccr5 in both Tet1- and Tet2-, but not Tet3-ablated cells. By contrast to the consistent loss of ectodermal marker expression detected in 1% O_2 cultured mESCs (Figure 3:11), Tet1-, 2- or 3-ablation largely trended to increase expression. Generally Tet3 silencing appeared to have a more pronounced effect than that of Tet1 and 2, resulting in marked induction of Fabp7[364], Dcx, Gbx2[365] and Tyr[366]. However, for genes

such as Hes5[367], Prom1[368], Neurog2[369], Galc[370] and Gfap[371], ablation of either Tet induced minimal to no change in expression, compared to shRNA control.

Overall the changes in expression of lineage-specific genes were highly variable, and therefore failed show a consistent trend towards (or away from) a specific germ layer. Thus, ablation of any one of the Tet enzymes failed to recapitulate the skewed differentiation of lineage markers observed in 1% O_2, compared to atmospheric O_2. Despite this, Tet1, Tet2 or Tet3 silencing did mimic the increased pluripotency marker expression that occurred with 1% O_2.

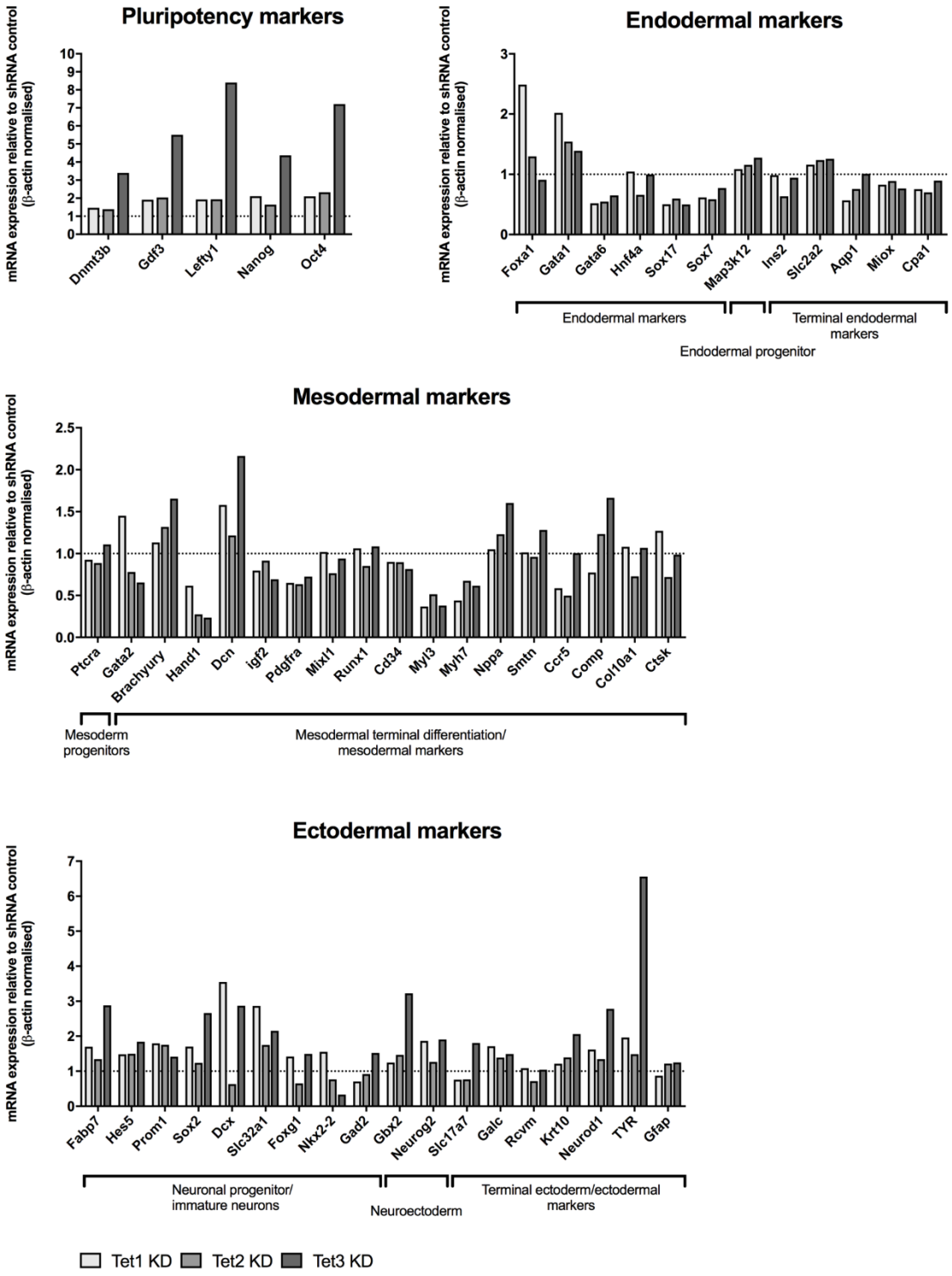


Figure 3:30 Tet ablation induces pluripotency marker expression. A QPCR array from pooled triplicate samples of day 7-differentiated Tet1, Tet2 and Tet3 ablated mESCs, showing expression of pluripotency, endodermal, mesodermal and ectodermal markers. Data are expressed as a fold change relative to shRNA control day 7-differentiated mESCs, depicted by the line at 1. KD: knockdown.

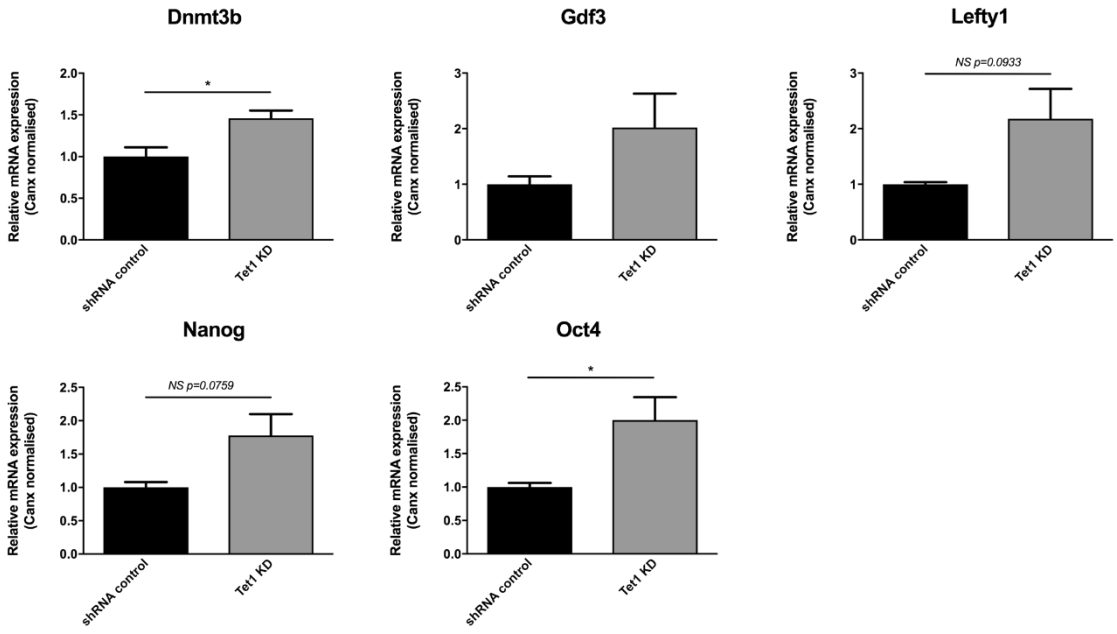


Figure 3:31 Validation of pluripotency marker expression in Tet1 silenced day 7-differentiated mESCs. Expression of Dnmt3b, Gdf3, Lefty1, Nanog and Oct4 in Tet1 ablated day 7-differentiated mESCs relative to shRNA control cells. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * p<0.05. KD: knockdown.

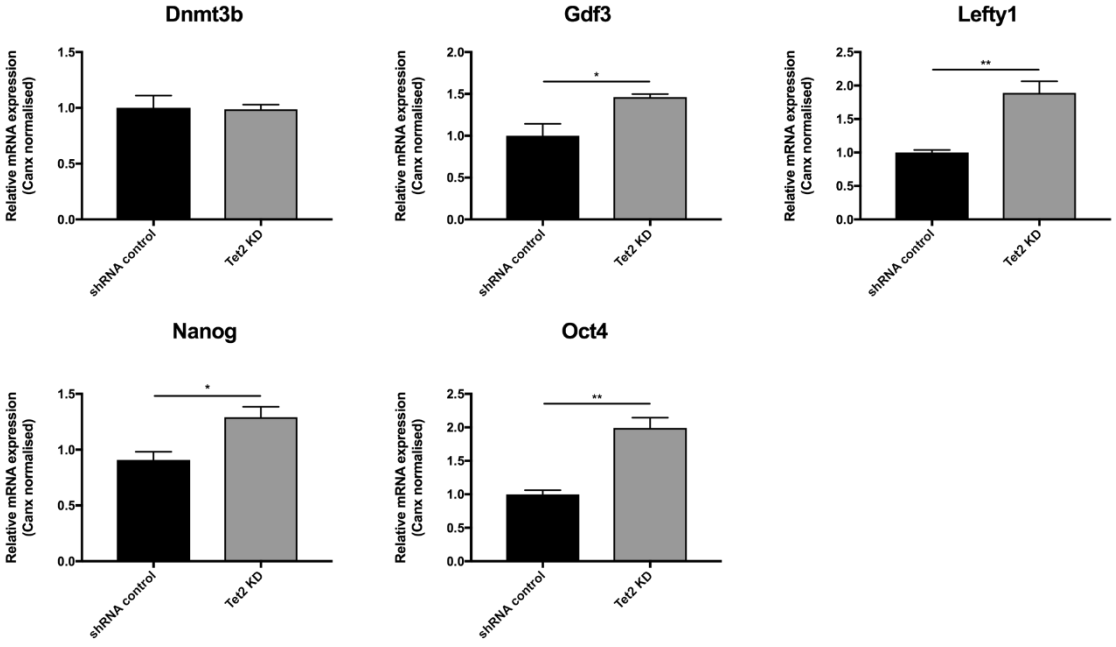


Figure 3:32 Validation of pluripotency marker expression in Tet2 silenced day 7-differentiated mESCs. Expression of Dnmt3b, Gdf3, Lefty1, Nanog and Oct4 in Tet2 ablated day 7-differentiated mESCs relative to shRNA control cells. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * p<0.05, ** p<0.01. KD: knockdown.

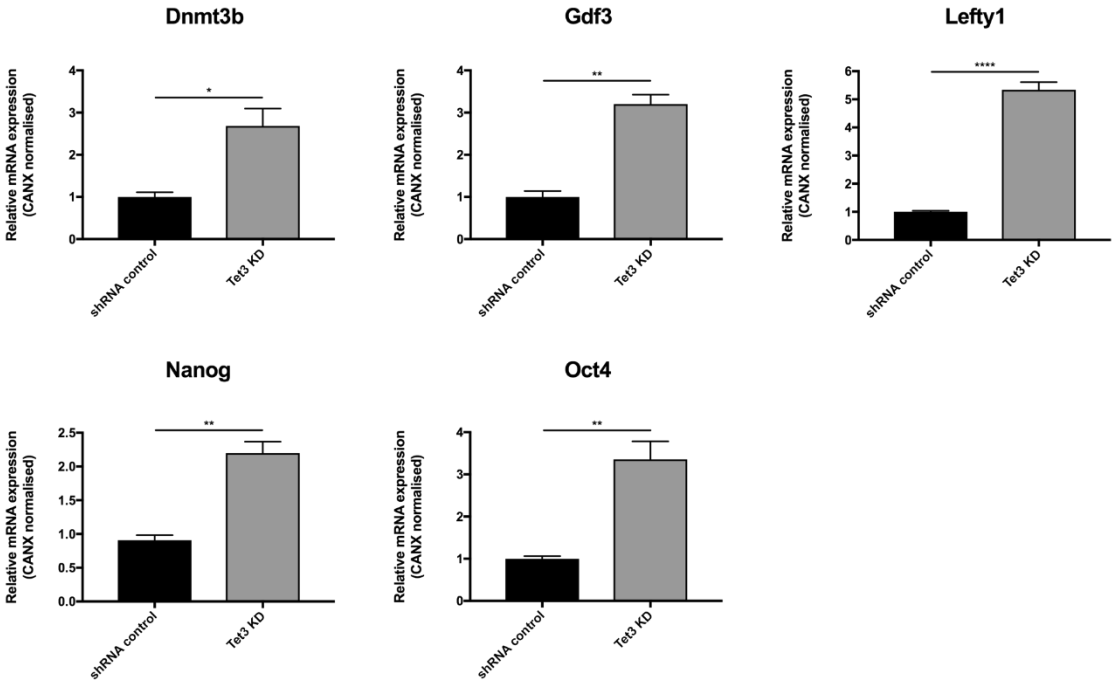


Figure 3:33 Validation of pluripotency marker expression in Tet3 silenced day 7-differentiated mESCs. Expression of Dnmt3b, Gdf3, Lefty1, Nanog and Oct4 in Tet3 ablated day 7-differentiated mESCs relative to shRNA control cells. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$. KD: knockdown

3.3.11 Microarray data revealed differential gene regulation by Tet1, Tet2 and Tet3 when overexpressed in P19 cells

To identify further the differential functional role of each Tet paralogue during cellular differentiation, Tet1, Tet2, and Tet3 were each overexpressed in P19 cells and induced to differentiate unbiasedly into EB-like aggregates for 2 days. Triplicate samples were pooled for microarray analysis (conducted by AROS Applied Biotechnology). Firstly, successful overexpression of each Tet was confirmed through QPCR (Figure 3:34). Tet1 showed the greatest induction, over 2000-fold, whereas Tet2 and Tet3 showed similar increases above pcDNA controls of 300- and 250-fold respectively.

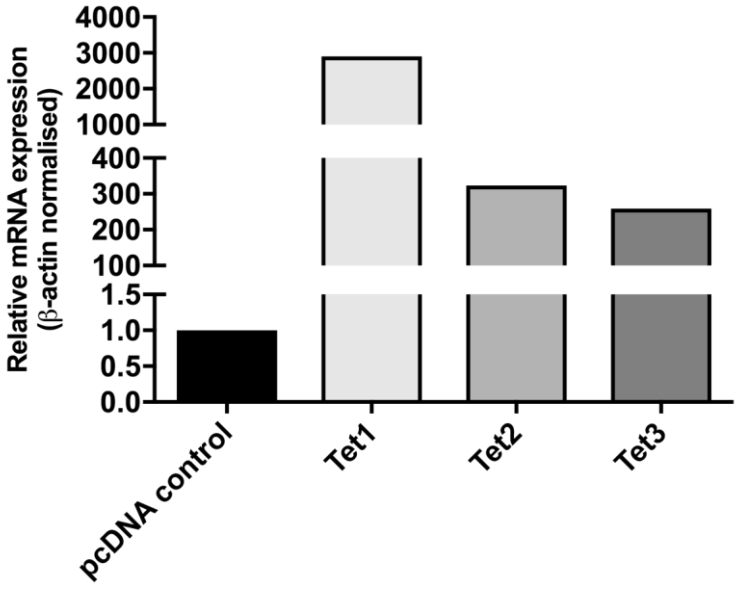


Figure 3:34 Overexpression of Tet1, Tet2 and Tet3 in undifferentiated P19 cells. pcDNA control or Tet1-, Tet2- and Tet3- overexpressing plasmids were transfected into undifferentiated P19 cells. Succesful overexpression was confirmed by QPCR analysis 48 hrs post-transfection, and data is expressed relative to pcDNA.

Microarray analysis revealed that Tet1, 2, or 3 overexpression induced differential expression changes, compared to pcDNA control transfected cells, with the majority of changes identified to be in fact non-coding RNA transcripts. The heat map, complete with hierarchical clustering (Figure 3:35A) demonstrates clear differences in expression patterns compared to pcDNA control transfected cells (red colours indicating increased expression and green decreased expression). This is confirmed by the Venn diagrams (Figure 3:35B) depicting the transcripts with a greater than 2-fold increase (shown in red) and greater than a 2-fold decrease (shown in green), compared to expression in pcDNA control transfected cells. It can be noted that some expression changes are consistent among all or 2 of the Tet paralogues. However, it is evident that each Tet regulates different transcripts suggesting they have independent functional roles during cellular differentiation. Perhaps unexpectedly only a small number of coding genes were identified, potentially owing to the short time point after the onset of differentiation at which cells were harvested. Despite this, intriguingly, the Tet enzymes did affect expression of short non-coding RNAs named miRNAs, which are responsible for post-transcriptional gene regulation. Table 6 shows the fold change for miRNAs that displayed a greater or less than 2-fold change relative to pcDNA transfected control cells. Some of these miRNA changes were specific for each Tet paralogue, such as miR-6344 induction with Tet3, while other miRNA expression was affected by all Tet paralogues, for example miR15b. Interestingly, Tets could also induce different effects upon the same miRNA. For instance, miR-378c expression was induced by approximately 9 and 12-fold with Tet1 and Tet2 overexpression respectively, but reduced around 3-fold with Tet3 overexpression. In light of this data, it may therefore be suggested that Tets contribute to gene regulation at the early stages of cellular differentiation indirectly through interacting with miRNA networks (see Discussion section 3.4.4).

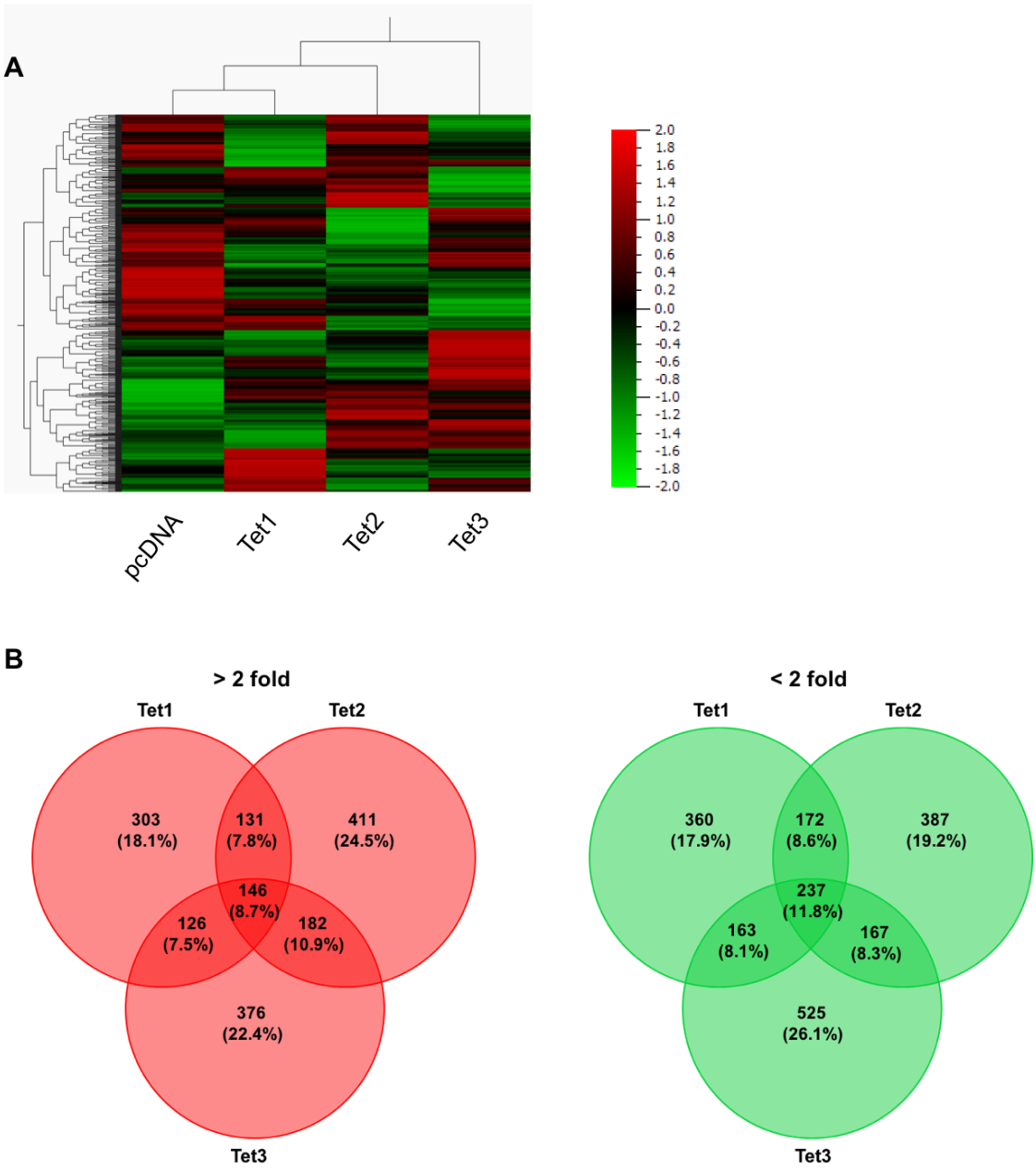


Figure 3:35 Microarray expression analysis revealed differentiated P19 cells overexpressing Tet1, Tet2 and Tet3 displayed differential transcript regulation. (A) Heat map with hierarchical clustering showing the differential expression of transcripts between pcDNA control and Tet1, Tet2, Tet3 overexpressing day 2-differentiated P19 cells. (B) Venn diagrams showing the number of uniquely and commonly expressed transcripts, which displayed a greater or less than 2-fold expression change, relative to pcDNA control, between Tet1, Tet2 and Tet3 overexpressing P19 cells.

miRNA	Tet1	Tet2	Tet3
miR-6344			2.10
miR-669m-1			2.23
miR-3098			3.73
miR-15b	2.92	2.97	0.27
miR-378c	8.79	12.48	0.35
miR-669a-1	5.14		
miR-669b	0.49	6.91	4.35
miR-669e	0.43		
miR-669j		2.16	
miR-669k		2.46	
miR-302c			0.44
miR-302a	2.22		0.32
miR-367			0.46
miR-3094	2.54	5.14	0.28
miR-467e	2.04	3.35	
miR-466l		2.23	

Table 6 Differential miRNA expression between Tet1, Tet2 and Tet3 overexpressing P19 cells. From microarray analysis miRNAs were found to be differentially regulated between Tet1, Tet2 and Tet3 overexpressing day 2-differentiated P19 cells. Here, the miRNAs, complete with their corresponding fold change values, relative to pcDNA control cells, are listed. Note only those miRNAs that demonstrated a greater or less than a 2-fold change in expression are shown.

3.4 *Discussion*

3.4.1 O_2 is a regulator of cellular differentiation

To assess the ability of O_2 to function as an important signalling molecule in differentiation 2 distinct methods were investigated. Firstly, differentiating P19 cells were treated with the HIF activator DMOG, the activity of which was confirmed through HIF-1α protein induction. It should be noted that this chemical mimetic of hypoxia is not a specific inducer of HIF, it will inhibit all 2-OGDOs and therefore recapitulate other HIF-independent O_2 signalling pathways. Unbiased differentiation of P19 cells yielded cells expressing lineage markers from all 3 germ layers, yet upon DMOG treatment, or pre-treatment, cellular specification could be skewed. Notably, the most consistent effects were seen on ectodermal marker expression, whereby all tested markers showed a decrease (or remain unchanged) after DMOG manipulation. Although not directly comparable, in a separate study the differentiation of P19 cells towards a neural fate demonstrated an increased number of dopaminergic neurons, when exposed to 3% O_2, compared to atmospheric controls[372]. This further highlights the ability of O_2 to regulate P19 cell differentiation, and is also in agreement with data shown here, suggesting a significant affect upon the ectodermal lineage.

The effects of DMOG provide a proof-of-principle concept that activation of O_2-sensitive signalling can induce changes in cellular specification. However, the use of chemical compounds is unlikely to fully recapitulate physiological mechanisms and may have additional off target effects. For example, from these studies, exposure of P19 cells to 1 mM DMOG for greater than 24 hrs, induced cell death. In addition, chemical mimetics may in fact resemble anoxia through complete inhibition of 2-OGDOs, whereas physiologically (potentially due to distinct K_ms) these enzymes likely display different sensitivities to [O_2]. A study directly comparing chemically-induced hypoxia with 5% O_2, on the secretion of pro-angiogenic factors from adipocytes, revealed that despite similarities, chemically mimicked hypoxia uniquely induced expression of matrix metalloproteinases, which was speculated to be caused by activation of different intracellular pathways[373].

To better mimic the physiological environment of embryogenesis, a ‘hypoxic chamber’ was used that could be controlled to a desired [O_2]. 1% O_2 was selected for differentiation studies to be in line with the range of O_2 tensions present in the uterus[71] and the stem cell niche environment[81]. Additionally, owing to the short diffusion distance of O_2, the formation of EBs under low [O_2] may simulate physiologically relevant O_2 gradients,

further mimicking early embryogenesis. Also, for future studies mESCs were used as a model of cellular differentiation. The cells used were an R1 mESC line, established from a 3.5-day blastocyst and therefore provide a more physiological relevant model of investigating cellular differentiation than P19 cells, which are isolated from teratocarcinomas[328,329].

Unbiased differentiation of mESCs, by contrast to P19 cells, resulted in a time-dependent decrease of both Oct4 and Nanog expression, suggesting the loss of pluripotency and thus onset of differentiation. QPCR analysis verified from the heterogeneous EB cell population the expression of key germ layer markers. Differentiation of mESCs exposed to 1% O_2 altered the pattern of differentiation significantly. Consistent with the literature[69], 1% O_2 promoted the undifferentiated state by an increase in an array of pluripotent markers. This may be mediated at least in part through HIFs, as the expression of Oct4 and Nanog has previously been shown to be regulated by HIF-2α[113,114,115]. Furthermore, HIF-1α accumulation under 1% O_2 has been shown to associate with the Notch intracellular domain, promoting its stability to activate Notch downstream targets Hes and Hey[374]. These can signal to regulate cell fate commitment[374], and Hes1 specifically has been shown to delay the exit from the pluripotent state by activation of Stat3[375] (a component of the LIF signalling pathway). Moreover, HIF-1α is also implicated in the activation of the Wnt/β-catenin signalling pathway[376,377], well known to promote pluripotency and self-renewal[378]. Thus overall, current evidence suggests that low O_2 tensions support the maintenance of a naïve pluripotent state at least in part *via* HIF signalling.

Induction of a selection of mesodermal markers under 1% O_2, suggested the mesodermal fate could be promoted under low O_2 tensions. Previously, expression of exogenous HIF-1α in mESCs has been shown to enhance cardiac specification through an increase in the beating phenotype of EBs *via* induction of both early cardiac markers Gata4 and Nkx2.5 and later markers α-myosin heavy chain 6 (Myh6) and β-myosin heavy chain 7 (Myh7)[119]. However, it should be noted that the expression of myosin-related genes from these studies were unaffected or decreased by 1% O_2 at this tested time point (Figure 3:11). It has also previously been demonstrated that the timing and duration of low O_2 tensions can impact upon cardiotypic differentiation. It has been demonstrated that exposure of 3% O_2 for 24 hrs, followed by normoxic (21% O_2) culture for a remaining 13 days, abolished the beating phenotype of iPSC derived cardiomyocytes[376]. However, this O_2 preconditioning was shown to enhance expression of the early mesodermal marker Brachyury[376], which was consistent with findings from DMOG treated P19 cells (Figure 3:7) and day 7-differentiated mESCs exposed to 1% O_2

(Figure 3:11). Thus, this supports the notion that [O_2] can regulate mesodermal differentiation.

mESCs showed consistent reduction of ectodermal marker expression when differentiated in 1% O_2 (Figure 3:15), similar to the effects of DMOG treatment in differentiating P19 cells (Figure 3:8). Specifically, 1% O_2 inhibited cells to acquire the neuronal lineage, demonstrated through abolishment of Dcx expression, a key gene in neurogenesis[379]. Previous studies have reported low O_2 tensions can promote neural differentiation[380,381,382], however one finding has demonstrated that O_2 can have differential effects dependent on the time course of neural development[383]. Specifically, it was shown that early neural differentiation of mESCs was inhibited in 5% O_2, yet later neural differentiation into neurospheres was promoted at this O_2 tension[383]. Thus, overall O_2 signalling is a critical factor for regulating cellular specification, consistent with its potential function as a morphogen during early embryogenesis.

As ESCs differentiate to become specialised cells they acquire distinct cellular identities with different phenotypes and functions, which are characterised by unique gene expression patterns. The determination of cell fate is dependent upon levels of chromatin compaction, which is dynamically adjusted through epigenetic modifications to regulate gene expression[384]. The structure of chromatin is itself liable to changes in O_2 tension[272,385], and therefore this led to the hypothesis that Tet enzymes owing to their 2-OGDO may in part mediate the O_2-dependent regulation of mESC function *via* alterations to DNA methylation dynamics.

3.4.2 Tet enzymes are potential O_2 sensors

For Tet enzymes to be considered as potential O_2 sensors during embryonic development they must of course be expressed. mRNA expression data from undifferentiated mESCs corroborates with previous data showing Tet1 to be the most abundant paralogue, followed by Tet2 and Tet3[386,252]. The presence of these enzymes, and their dependence on O_2 for catalytic activity, provides a plausible basis to test their ability to act as O_2 sensors. Firstly, assessment of the functionality of each Tet under atmospheric [O_2] revealed that Tet2 showed the greatest induction of 5-hmC, compared to Tet1 and Tet3 (Figure 3:22). Western blot analysis confirmed this did not occur because of different levels of ectopic expression, suggesting Tet2 has a higher level of catalytic activity in baseline atmospheric O_2 conditions. The reason for this is not clear but this may be due to the lack of the DNA-binding CXXC domain, which is present in Tet1 and Tet3[173]. This potentially may mean that Tet2 has less specificity, reflective of

its more promiscuous nature.

Exposure of Tet overexpressing HEK cells to decreasing O_2 concentrations, considered relevant within the physiological range of mammalian embryogenesis[69,81,71], revealed Tet1 as the most likely candidate as an O_2 sensor, owing to inhibition of its catalytic activity at O_2 tensions less than 3% (Figure 3:23 and Figure 3:24). However, it should also be noted that Tet3 activity was unexpectedly elevated at 3% O_2 and therefore may also be considered a sensor at this [O_2]. Tet2 showed inhibition only at the lowest tested [O_2], which may be attributed to a high affinity for O_2 and thus a low K_m. The use of DMOG (that may resemble an anoxic state), which is presumed to give maximal inhibition, only induced a 30% reduction in activity (compared to 70% inhibition of Tet1 activity), in line with effect of 0.5% O_2, thereby suggesting Tet2 is also resistant to chemical inhibition. Other reports have suggested, through K_m calculations, that Tet1 and Tet2 have similar sensitivities to O_2[303,309]. However, it can be argued that the recombinant systems used in such studies are not as physiological as the cellular model used here.

Analysis of TCA cycle metabolites in 3 and 1% O_2 exposed HEK cells revealed that differential succinate levels may account for elevated Tet3 activity at 3% O_2 (Figure 3:25). The activation of the anaerobic glycolytic pathway would suggest that TCA metabolites would decrease, however as in agreement with the literature the succinate levels were found to increase under low O_2 tensions[319,320], suggesting that O_2 induces secondary effects that may in part underlie the regulation of Tet-specific functions.

For the first time this data shows the differential regulation of activity of all 3 Tet enzymes by O_2. Thus, it can be postulated that their spatial distribution within the early embryo would enable exposure to O_2 gradients, resulting in differential activity, which may in turn determine activation of cellular transcriptional pathways regulating cell fate.

3.4.3 Tet proteins are implicated in the pluripotency regulatory network

To investigate whether Tets can have a direct function on cell specification and resemble the skewing of mESC differentiation induced by 1% O_2, genetically ablated mESCs were inducted to differentiate for 7 days and expression of developmental and lineage markers assessed (Figure 3:30) (as per for 1% O_2 studies).

Ablation of each Tet induced expression of all tested pluripotency markers (Figure 3:31, Figure 3:32 and Figure 3:33), consistent with the effects of 1% O_2, suggesting that Tet enzymes are in part responsible for mediating this response. The implication of Tets in

the pluripotency regulatory circuit is well reported, since the enrichment of 5-hmC has been noted at binding sites of pluripotency related TFs[185], including Oct4 and Nanog[387]. Further analysis has revealed that Tet1 and Tet2 are downstream targets of Oct4[252], and genome-wide profiling of Tet1 distribution in mESCs correlated with the regulation of pluripotency and differentiation-related gene expression[388]. This may therefore suggest that Tets have a functional role in regulating the balance between pluripotency and differentiation events.

3 independent groups using ChIP-seq revealed Tet1 preferentially binds at gene promoters *via* CpGs[389,183,194], and is associated with 5-hmC[390], supporting a hypomethylated state. Tet1 binds to actively transcribed pluripotency-associated factors, including Nanog[390], suggesting a role of Tet-mediated hydroxylation in promoting a transcriptionally active state enabling recruitment of TFs and basal transcriptional machinery[388]. To support this notion, Tet1 depletion resulted in a decrease in Nanog expression, which correlated with 5-mC on its promoter[386]. In a separate study, Tet1 and Tet2 ablated mESCs resulted in increased promoter methylation and consequently decreased expression of the pluripotency-related genes Esrrb, Klf2, Tcl1, Zfp42, Dppa3 and Ecat1 (albeit Oct4 and Nanog were unaffected)[186]. A more recent report has confirmed the physical association of Tet1 and Tet2 with Nanog, and interestingly were shown to be functionally redundant in enhancing the reprogramming and establishment of naïve pluripotency[391]. However, there are discrepancies in Tet silencing studies, Tet1-depleted cells have been shown to not downregulate Nanog and Oct4 expression[252,389]. Furthermore, Tet1-depeleted mESCs stained positive for alkaline phosphatase (high expression of alkaline phosphatase in ESCs provides a measure of pluripotency) and retained expression of Oct4 and Nanog[251]. The authors also took additional steps to confirm the pluripotency and differentiation capabilities of Tet1 depleted ESCs by a tetraploid complementation assay (a tetraploid embryo will only develop if combined with functional diploid ESCs)[251].

Inconsistencies in ESC studies are common due to the ESC line used, culture conditions, method of Tet silencing and Tet compensatory effects. In addition, transcriptional control by Tet1 has added complexity due to an apparent dual function, not only in the activation of pluripotency genes as mentioned, but also repression of Polycomb marked developmental genes[194] (see Chapter 4: Results 2). Perhaps controversially, the data collected here appears to demonstrate that Tet proteins are acting to repress pluripotency genes, however this expression analysis was performed on day 7-differentiated samples, not at the undifferentiated level as investigated by many groups. It can be noted the induction of pluripotency genes by Tet1 and Tet2 silencing does not

fully mimic the extent of increased expression seen with 1% O_2, which proposes the ability of other O_2 sensing pathways, such as HIFs[114], to also be implicated.

Intriguingly, despite a lack of literature support, Tet3-silencing here induced pluripotency marker expression and to a greater extent than Tet1 and Tet2 ablation. However, a potential indication for the involvement of Tet3 in pluripotency regulation can be inferred from a study in which Tet1 and Tet2 double knockout mESCs exhibited a 2-fold induction of Tet3 mRNA and remained pluripotent[248]. A further insight was gained from somatic cell nuclear transfer experiments, which utilised somatic cells containing an Oct4-EGFP transgene[392]. Compared with embryos derived from wild-type oocytes, Tet3-null oocytes displayed reduced EGFP signal, consistent with decreased Oct4 mRNA levels[392].

One possible explanation for Tet3 ablation having the most dominant effect on pluripotent marker expression in the studies shown here is that during differentiation, Tet3 expression increases (while Tet1 and Tet2 decrease)[386] (see Figure 4:5). Therefore, at the tested day 7-time point Tet3 is likely to have a more intrinsic role in cellular differentiation and development. However, silencing of Tet3 for 7 days resulted in a compensatory increase in Tet1 expression (data not shown), which may have contributed to the increase in pluripotent genes, but also infers a possible regulatory link between these 2 Tet paralogues (see Chapter 6: General Discussion). Overall, these data suggest Tet proteins may be intrinsic to the balance of pluripotency and lineage commitment[393].

3.4.4 Tet proteins and miRNAs: a novel gene regulatory mechanism

To gain further insights into the differential role of Tet proteins in cellular differentiation, each Tet was overexpressed into P19 cells and induced to differentiate into EBs. Microarray analysis surprisingly revealed Tet overexpressing cells induced differential expression of miRNAs, rather than lineage-related genes. This unexpected result may have occurred because cells were harvested after a short period (2 days) of differentiation, which was conducted to ensure the effects of transient Tet overexpression were not lost. Intriguingly however, identification of a transcriptional regulatory network between Tets and miRNAs may provide further understanding of gene regulation at the onset of cellular differentiation.

miRNAs regulate gene expression at the post-transcriptional level *via* targeting mainly the 3'-untranslated regions (3'UTRs) of mRNAs[394]. In the biogenesis pathway miRNAs are transcribed by RNA polymerase II to form long stem-loop containing primary-

miRNAs (pri-miRNAs)[394]. Here within the nucleus, a ribonuclease III (RNase III), named Drosha and its associated RNA-binding protein, DiGeorge syndrome critical region 8 (Dgcr8) crops pri-miRNAs into hairpin-shaped precursor-miRNAs (pre-miRNAs) of ~70 nucleotides in length[394]. Pre-miRNAs are transported to the cytoplasm *via* exportin-5 where Dicer cleavage forms a ~22 nucleotide mature miRNA duplex[394]. The mature miRNA products can be assigned 5p or 3p to represent the arm of the pre-miRNA they formed from, in the canonical model one strand of the duplex is biologically active and the other (passenger or * strand) is thought to be inactive and degraded[395], although this concept has been challenged[396]. The guide strand becomes incorporated into the Ago2 protein containing RISC, facilitating the stabilisation of miRNA:mRNA interactions. The guide strand binds to its target site in the 3'UTR of the mRNA primarily through perfect pairing of a 2-7 nucleotide sequence at its 5' end, known as the seed region[395]. This subsequently results in the repression of gene expression *via* translational inhibition or mRNA degradation.

The regulation of Tets post-transcriptionally by miRNAs has been reported previously in the cancer field. For example, Tet2 has been shown to be under extensive miRNA regulation (>30 miRNAs) resulting in diminished expression and subsequent 5-hmC loss[397]. However, there is also evidence for such interplay in developmental events. For instance miR-26a has been demonstrated to act to decrease mRNA levels of all Tet paralogues and is implicated in neuronal[398] and pancreatic cell differentiation[399]. In addition, targeting of Tet1 by miR-29[400] is implicated in repressing Tet1 expression after EB formation and promotion of mesoendodermal lineage formation[401].

miRNAs are an integral part of the pluripotent regulatory network. Initial studies in mESCs depleted of Dicer, an intrinsic component of miRNA biogenesis, failed to differentiate and form teratomas[402]. Additionally, Dicer mutant embryos have been shown to lack Oct4 staining[403]. Clusters of miRNAs (a set of 2 or more miRNA genes that are transcribed from physically adjacent genes[404]) have been shown as key regulators of pluripotency and differentiation events[405]. Specifically, the promoter region of the miR-302 gene, which encodes a cluster of 8 miRNAs (miR-302b*-b-c*-c-a*-a-d-367), is bound by Oct4, Sox2 and Nanog[406,407]. miR-302 is highly expressed in undifferentiated P19 cells[408], so perhaps it was unsurprising that it was detected from the microarray shown here after a short period of P19 cell differentiation (Table 6).

The miR-302 family being highly conserved and abundantly expressed in undifferentiated ESCs[409], has had invested interest in its somatic cell reprogramming capabilities to iPSCs[410,411,412,413]. Interestingly, one proposed mechanism of reprogramming is mediated through miR-302 targeting of epigenetic factors to reset

global DNA demethylation patterns (in addition to other chromatin modifications)[413]. For example, miR-302 silences MBD2 (important for zygotic demethylation, despite not a fully understood mechanisms) as well as Dnmt1, meaning DNA methylation can't be replicated, subsequently leading to global demethylation[410,412,414]. Furthermore, latest mechanistic findings linking miR-302 to pluripotency revealed high endogenous expression of miR-302 maintained Oct4 levels *via* suppression of AKT1, and reversely showed, through the use of a miR-302 antagomir the induction of differentiation through teratoma formation[415].

Besides one study demonstrating Tet2 mRNA is repressed by the miR-302/367 cluster[416], there is limited evidence for this miRNA cluster in Tet regulation. From this data (Table 6), miR-302a/c was found to be differentially regulated by Tet1 and Tet3. miR-302a was found to be upregulated in Tet1 overexpression, supporting the notion that Tet1 is an intrinsic regulator of pluripotency, whereas miR-302a, miR-302c and miR-367 were down regulated with Tet3 overexpression. Genetic ablation of the miR-302 cluster *in vivo* resulted in early developmental defects in the neural tube, residing as embryonic lethal, suggesting a role of miR-302 in neurulation events[417,418]. Thus, the Tet3-mediated downregulation of miR-302 demonstrated here may link Tet3 to neural development. In support of this neurogenesis role, miR-15b was also shown to be downregulated with Tet3 overexpression (Table 6). miR-15b has been shown to promote neuronal differentiation, and directly target the 3' UTR of Tet3 leading to its repression and subsequent decreased 5-hmC levels[419].

It should be considered that overexpression of Tets are likely to affect many miRNA circuits, which may result in changes in miRNA expression that are not in fact direct Tet targets. However, investigation into the putative miRNA binding sites within the 3'UTR of the Tet proteins, using the miR-walk data base (Available at: http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/) revealed that Tet1, Tet2 and Tet3 were predicted to be bound by miR-320a/c and miR-15b. The ability of Tet overexpressing P19 cells to alter expression of these miRNAs (Table 6) in part confirms the presence of these miRNA/Tet regulatory networks. However, these miRNA expression changes need to be further validated by QPCR. A final speculative point is that as miRNA expression can be regulated by DNA methylation[420] and $[O_2]$[421], it may thus be plausible to suggest that O_2-dependent regulation of Tet activity may directly impact miRNA expression to orchestrate a complex gene regulatory network during embryogenesis.

3.4.5 Conclusion

To conclude, changes in [O_2] were shown here to affect the patterning of mESC cellular differentiation, consistent with the notion of O_2 acting as a morphogen during early embryogenesis. A more hypoxic environment (1% O_2), compared to atmospheric O_2, delayed the onset of mESC differentiation through an increase in pluripotency marker expression and predominantly skewed differentiation away from the neuroectoderm lineage. Tet enzymes were investigated to determine whether at least in part they could mediate the skewing effects of low O_2 upon cellular differentiation and thus act as O_2 sensors. For their sensor capabilities to be true, the presence of Tet enzymes in ESCs was firstly confirmed. Secondly, the activities of the Tet enzymes were shown to display differential regulation by physiologically relevant O_2 concentrations. Thirdly, functionally, genetic ablation of each Tet was shown to induce pluripotency marker expression during mESC differentiation (consistent with the effects of 1% O_2) and *via* Tet overexpression studies were each shown to differentially regulate miRNA expression. However, it should be noted that differentiation of Tet ablated mESCs failed to recapitulate the skewing of lineage markers seen by 1% O_2, alluding to the involvement of additional O_2-dependent developmental pathways and/or Tet compensatory effects (see Chapter 6: General Discussion).

Chapter 4: Results 2; *5-hmC in Cellular Differentiation*

4.1 *Introduction*

4.1.1 Dynamic expression of 5-hmC in differentiating ESCs

Implantation of the blastocyst to the uterine wall results in depletion of ESCs as they start to differentiate into more specialised cell types. Attempts to recapitulate early embryonic events by the differentiation of ESCs *in vitro* revealed the loss of 5-hmC, compared to undifferentiated cells, after 5 days of EB formation[168]. Additionally, it was independently shown, *via* a glucosylation assay, that 5-hmC was lost after 4 days of EB formation in 2 separate mESC lines, compared to levels in undifferentiated cells[422]. Intriguingly, 5-hmC in differentiating EBs were noted to increase from day 4 to day 8, but levels remained over 50% lower than in undifferentiated cells[422]. Alternatively, by a mass-spectrometry approach, it was shown 5-hmC in day 1-differentiated cells was elevated, compared to undifferentiated cells, and then subsequently gradually decreased to day 7[423]. 5-hmC levels were then see to increase to levels equivalent to undifferentiated cells by day 10[423]. However, it should be considered the calculation of 5-hmC levels was performed by normalisation to the combined sum of 5-mC and 5-hmC, not total C (C, 5-mC and 5-hmC) and thus will not account for changes in C that may arise from proliferation. In comparison, 5-mC levels showed a gradual increase at the onset of differentiation, but plateaued by day 3[423]. These distinct profiles may therefore suggest different functional roles of 5-hmC and 5-mC in cell specification. However, the exact dynamic profile of 5-hmC during ESC differentiation requires further elucidation with a more valid method of quantitation.

4.1.2 Activation of Tet enzymes by Ascorbic acid may control 5-hmC generation and distribution during embryogenesis

The importance of a dietary intake of AA, commonly known as vitamin C, was brought about through the ability of oranges and lemons to treat scurvy in the early 18th century[424,425]. Scurvy symptoms, such as gum disease and poor wound healing are a result of defective collagen production[125,426]. The synthesis of hydroxyproline, a major component of collagen, occurs enzymatically by collagen prolyl-4-hydroxylase[427]. This enzyme belongs to the 2-OGDO family and requires AA as a cofactor for its activity[428]. It is believed that AA maintains 2-OGDOs in an active state, thus enabling continued cycling of the enzymes, by the reduction of Fe^{3+} to Fe^{2+} [125,427]. This is supported by an investigation into the effects of AA upon PHD inhibition induced by either O_2 deprivation, 2-OGDO competition, iron chelation or transition metal-induced toxicity[429]. AA had a

profound effect on preventing HIF-1α protein stabilisation specifically in response to iron competition[429] (the iron chelator desferrioxamine (DFO), Co^{2+} and Ni^{2+} all induced HIF-1α through poisoning of the hydroxylases *via* removal of enzyme-bound iron[125]). The importance of AA for 2-OGDO activity can also be depicted from K_m values, which for PHDs range from 140-180 μM[430], suggesting a high intracellular requirement and therefore susceptibility to AA loss[125].

The requirement of 2-OGDO's for AA suggests that Tet-mediated 5-hmC generation may be regulated in part through this co-factor. Indeed, it has previously been observed that AA can induce DNA demethylation[431,432,433] through enhanced Tet activity[434]. In mESCs it has been reported that the addition of AA induces not only 5-hmC but also the further oxidation products 5-fC and 5-caC, suggesting activation of the active DNA demethylation pathway[434]. Furthermore, this study demonstrated that Tet1- and Tet2-silenced mESCs failed to induce 5-mC oxidation in the presence of AA, an effect restored upon re-expression of Tet2[434]. This finding was closely mimicked by an additional study, which confirmed the effects of AA are Tet-dependent, and further demonstrated that AA induced widespread DNA demethylation at the promoters of germ-line genes, resulting in their subsequent activation[435]. Interestingly Tet1 was found to be enriched at promoters that gained 5-hmC with AA treatment[435], which may suggest Tet1 mediates the AA-induced 5-hmC at developmental genes. Additionally Tet1 was found to be implicated in somatic cell reprogramming in a AA-dependent fashion[436]. This finding may be due to the necessity of AA to induce hydroxylation and therefore maintain low levels of 5-mC, required for naïve pluripotent stem cells[437]. Overall, this suggests AA may function as a switch, dependent on exact concentrations, to determine Tet1 function.

This effect of AA was also found to be applicable to other cell types including MEFs, which demonstrated a time-dependent increase in 5-hmC with AA treatment, an effect lost by combined siRNA targeting of Tet1, Tet2 and Tet3. Moreover, AA is highly concentrated in the brain, especially during embryogenesis, and facilitates the differentiation of neural stem cells into dopaminergic neurons[438]. Mechanistically, enhanced transcription of dopaminergic neuronal genes was associated with AA-mediated induction of 5-hmC by Tet1 within promoter regions[439]. In addition, AA has also been shown to potentiate Tet2 and Tet3, which functionally has been demonstrated to increase stability of the Foxp3 TF, required for the development and function of regulatory T-cells[440].

Finally, it should be considered that unlike HIFs where AA can be substituted for the GSH reducing agent[430], Tet enzymes have an absolute requirement for AA

specifically[434,435,441]. Therefore, the availability of AA may have implications in health and disease through epigenetic control of the genome[441].

4.1.3 Tet proteins cross-talk with histone modifications to regulate gene transcription

Aside from Tet-mediated DNA demethylation as a means of gene-transcriptional regulation (see Chapter 1: General Introduction), Tet proteins co-operate with histone marks to control transcription of developmental genes[189].

Nucleosomal histones can be post-translationally modified by addition of the monosaccharide β-N-acetylglucosamine (GlcNAc) to the hydroxyl group of serine or threonine residues to form an O-linked β-N-acetylglucosamine (O-GlcNac) residue[442,443]. The process of O-GlcNacylation is mediated by OGT, which was unexpectedly found to interact with Tets[189]. OGT complexes with Tet1 and Tet2 in the nucleus of ESCs, and Tet1 specifically has been shown to co-localise with OGT at H3K4me3 positive promoters[190]. Intriguingly, Tet1 and OGT were also found to complex, with Hcf1. Hcf1 is a component of the H3K4 methyltransferase SET1/COMPASS complex[444] and is modified by OGT[192] to elevate H3K4me3 to induce transcriptional activation[189]. Furthermore, genetic ablation studies in HEK cells showed the association of Tet2 and Tet3 with OGT at H3K4me3 enriched promoters was essential for transcriptional activation[191]. Mechanistically it was shown in this model the Tet-OGT complex targeted Hcf1 to control levels of H3K4me3[191]. Overall, this suggests a functional connection between OGT, Tet proteins and the transcriptional activating H3K4me3 mark.

Another prominent mechanism of modifying chromatin is through the Polycomb group of proteins (PcG). PcG protein complexes are classified into 2 major families, Polycomb repressor complex 1 (Prc1), which has a E3 ubiquitin ligase activity for the monoubiquitylation of histone H2A, and Prc2, which has methyltransferase activity for the generation of H3K27me2/3[445]. Prc2 has a core trimeric protein complex containing the Ezh1/2 catalytic subunit responsible for depositing H3K27me2/3 marks[446,447], facilitating the recruitment of additional PcG proteins[448] to contribute to the transcriptional repressed chromatin state.

Cross-talk between 2 epigenetic silencing mechanisms, DNA methylation and Prc2-H3K27me2/3 is widely evident since the discovery of an interaction between Dnmts and the Ezh2 subunit[449]. In support of this, it was later demonstrated in DNA hypomethylated somatic cells that H3K27me3 and Prc2 was lost from target gene promoters, which

subsequently resulted in ectopic expression in some of these genes[450]. Alternatively it has been reported that antagonism exists between 5-mC and H3K27me2/3 at the murine imprinted Rasgfr1 locus, whereby either mark prevents placement of the other and are thus considered mutually exclusive[451]. Furthermore, unmethylated CG regions have been linked to the recruitment of the Prc2 complex[452,453] and perturbation of DNA methylation patterns resulted in redistribution H3K27me3[454]. Finally, in mESCs the loss of 5-mC through ablation of Dnmt activity resulted in the induction of H3K27me3 at sites of fully methylated regions in wild-type cells[455]. Therefore, in light of these data the exact relationship between 5-mC and Prc2 remains elusive, but further insight may be gained into this transcriptional regulatory complex through consideration of the role of DNA demethylation and thus the 5-hmC mark.

The current understanding of Tet involvement with PcGs is relatively unexplored, however it is documented that in mESCs Tet1 binds to CG rich promoters of Polycomb-repressed genes[194]. In a separate study, unique to undifferentiated mESCs, 5-hmC was found to co-localise with Prc2 and in fact Prc2 was found to be responsible for recruitment of Tet1 to H3K27me3 genomic regions[456]. However, it should also be noted that the reverse effect has been reported, such that Tet1 itself has been shown to facilitate Prc2 recruitment[194]. Functionally, in differentiated cells, it has been shown that Prc2 can function to decrease Tet1 expression *via* H3K27me3 deposition[457], which therefore may present a mechanism to silence the Tet1-dependent pluripotency network upon induction of differentiation.

The active (H3K4me3) and repressive (H3K27me3) histone marks can coexist at developmental genes within pluripotent ES cells, a situation referred to as bivalency[458]. This concept of bivalency, having 2 apparent functionally opposing histone marks together at a promoter, enables genes to be actively poised for rapid activation[193,459]. The presence of the H3K27me3 mark on lineage specific genes within pluripotent cells is necessary to prevent premature expression, as demonstrated by the increase in ectoderm differentiation upon disruption of the Prc2 complex[460]. Therefore, in response to developmental signalling cues promoters become resolved into monovalent chromatin structures of either H3K4me3 and H3K27me3, resulting in gene activation or silencing respectively[460]. Genome-wide profiling revealed promoters of genes such as developmental TFs in ESCs with a high CpG density carry this bivalent signature, but this is not always limited to pluripotent cells[461]. These initial *in vitro* studies reporting bivalency involve the maintenance of pluripotency in a simulated environment, and therefore their existence has been linked to the culture conditions[462]. However, this

bivalency phenomenon has been shown to occur *in vivo* using an adapted ChIP methodology in mouse Epi cells of the early post-implantation embryo[463] and in PGCs[464].

As described above, there is a clear link between Tets/5-hmC and these 2 histone methylation marks, but only recently it was made apparent that Tets are involved in the *de novo* establishment and maintenance of the H3K4me3/H3K27me3 bivalent domains[465]. Tet2 overexpression in HEK cells induced loci-specific bivalent domain formation at promoters of genes normally found hypermethylated and lacking these histone methylation marks[465]. Furthermore, depletion of Tet1/Tet2 in mESCs led to the loss of H3K27me3 and thus bivalency[465]. No effect upon H3K4me3 was evident likely due to binding of this histone mark to CpGs that are unmethylated[466]. Overall, this crosstalk between Tets and histone methylation highlights the ability to fine-tune gene transcription. It also may be plausible to speculate that O_2 is involved in the regulation between these epigenetic marks owing to the 2-OGDO domain, which is not only present in Tets but also within the JmjC-KDMs.

4.2 *Aims*

To determine the profiles of 5-mC and 5-hmC during a time course of mESC differentiation exposed to atmospheric and low [O_2]. In addition, the aim is to identify the Tet enzymes responsible for regulating 5-hmC during cellular differentiation and investigate any potential cross-talk with histone methylation marks.

4.3 *Results*

4.3.1 A transient burst of 5-hmC was detected during mESC differentiation

To begin to investigate the profile of DNA methylation marks during cellular differentiation, the global levels of 5-mC and 5-hmC over an 11-day time course of mESC differentiation were determined by dot-blot analysis (Figure 4:1). This revealed that 5-mC levels increased between days 1 and 3, and continued to be maintained at this expression level over the remaining differentiation time course. 5-hmC demonstrated a more dynamic profile, showing a significant burst at day 3, which subsequently decreased to levels equivalent to those of undifferentiated cells by day 11. This transient burst of 5-hmC was also confirmed by analysing the levels of 5-mC in the form of a 5-hmC:5-mC ratio at each of the tested time points. To verify these findings mass spectrometry was used to detect C, 5-mC and 5-hmC, enabling the amount of 5-hmC to be normalised to total C within each sample (C, 5-mC and 5-hmC) (Figure 4:2). This analytical approach reduced the variability within sample groups, compared to dot-blot analysis, and confirmed the expression profile of both methylation marks. Interestingly, the significant elevation of 5-hmC at day 1 was not preceded by increased 5-mC levels, which may infer that hydroxylation occurred at already existing methylated residues or that methylation and hydroxylation are closely coupled.

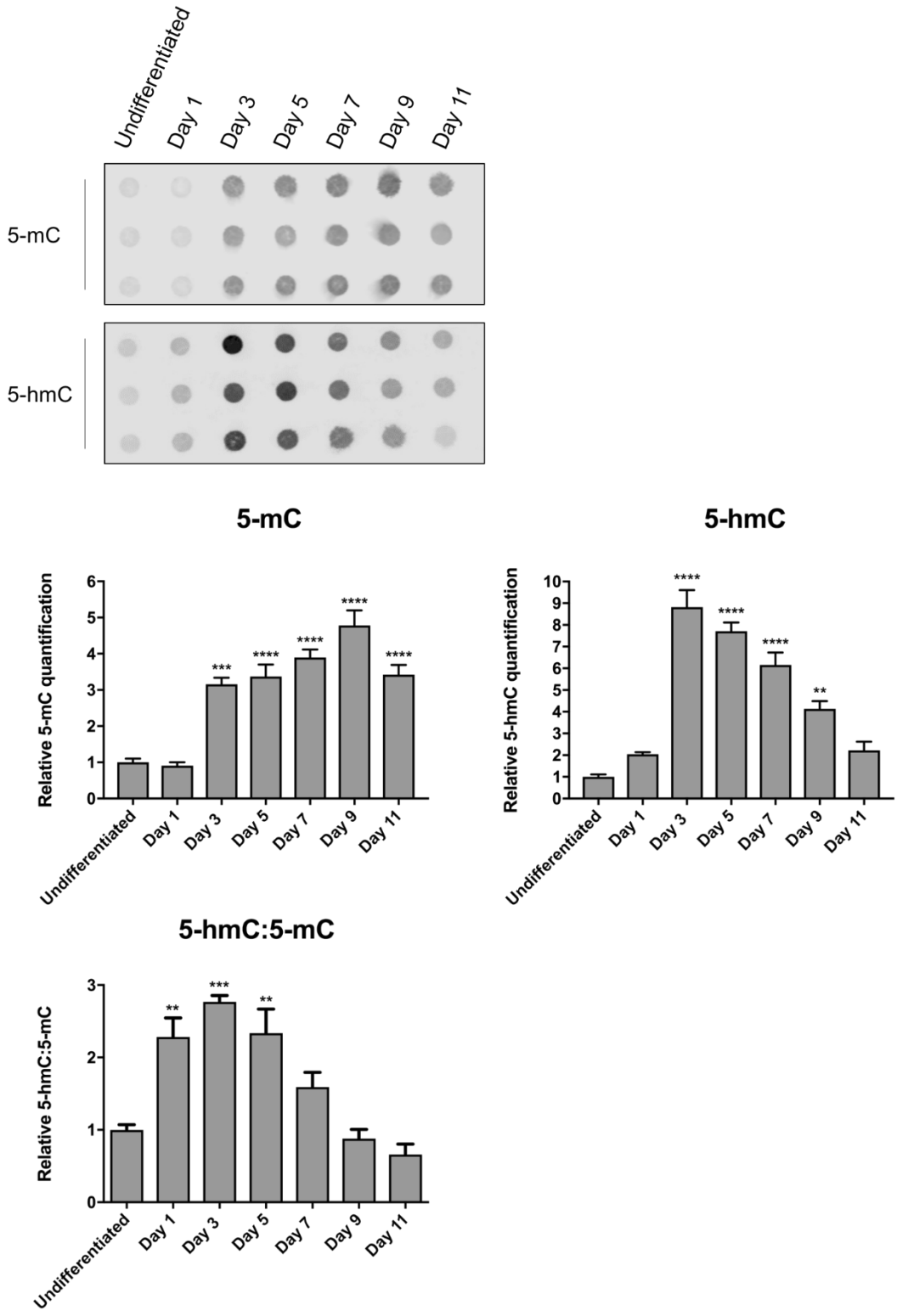


Figure 4:1 mESCs demonstrate a burst of 5-hmC at day 3 of cellular differentiation. mESCs were induced to differentiate over an 11 day time course and DNA harvested to assess 5-mC and 5-hmC levels by dot-blot analysis. Quantifcation of 5-mC, 5-hmC and the ratio of 5-hmC:5-mC is shown relative to undifferentiated mESCs. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. ** p<0.01, *** p<0.001, **** p<0.0001

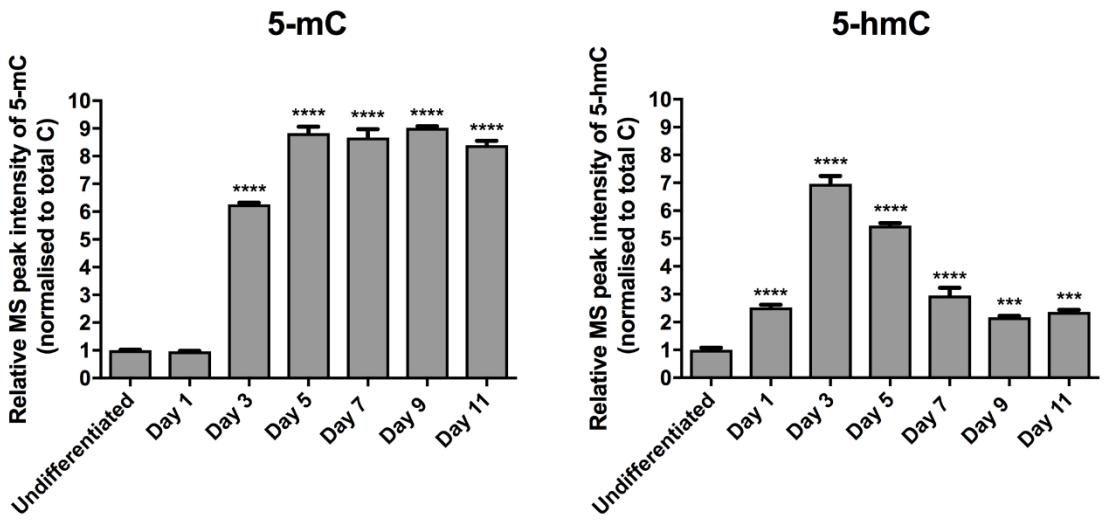


Figure 4:2 Mass-spectrometry analysis confirms a burst of hydroxylation in day 3-differentiated mESCs. mESCs were differentiated over an 11 day time course and DNA harvested for C, 5-mC and 5-hmC quantitation by mass-spectrometry analysis. Data are expressed as the mean ± SEM relative to undifferentiated cells. Data are normalised to total C and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. *** p<0.001, **** p<0.0001

4.3.2 Compared to atmospheric O_2, levels of 5-hmC during mESC differentiation are reduced in 1%, but not 3% O_2

To establish whether altering [O_2] can affect the dynamic 5-hmC profile during differentiation, mESCs were induced to differentiate over the same time course as shown in section 4.3.1, but in 3% and 1% O_2. Dot-blot analysis (Figure 4:3) revealed 3% O_2, compared to atmospheric O_2, had minimal effects on the 5-hmC profile, the only significant change being a rather unexpected increase in hydroxylation at day 7. By contrast, 1% O_2, compared to atmospheric O_2, consistently and significantly reduced 5-hmC levels at all time points, except for days 1 and 11.

Compared to dot-blot analysis, mass-spectrometry (Figure 4:4) reported additional significant differences in 5-hmC levels under 3% O_2, compared to atmospheric O_2. In light of these data it may be suggested that 3% O_2 acts to delay the burst of 5-hmC at day 3, and subsequently prevents its decrease until day 7. Furthermore, under 1% O_2 the blunting of the 5-hmC burst at day 3 was clearly evident, and additionally each differentiation time point showed significantly reduced levels of hydroxylation. Finally, it can be noted that neither [O_2] had a significant effect on levels of 5-hmC in undifferentiated mESCs.

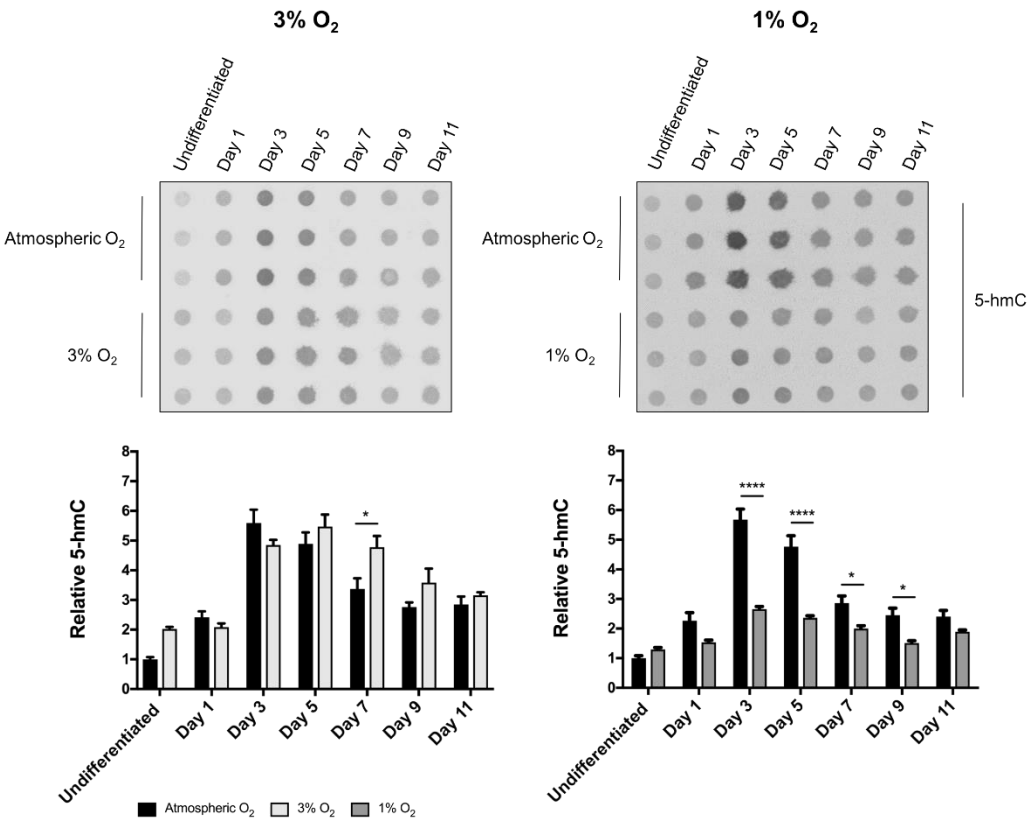


Figure 4:3 1% O_2, but not 3% O_2, reduces 5-hmC levels during mESC differentiation, compared to atmospheric O_2. mESCs were differentiated over an 11 day time course in atmospheric, 3% and 1% O_2 and levels of 5-mC and 5-hmC were measured by dot-blot analysis. Data are expressed as the mean ± SEM relative to undifferentiated mESCs exposed to atmospheric O_2, and analysed by a two-way ANOVA with Sidak's post-hoc analysis. * $p<0.05$, **** $P<0.0001$

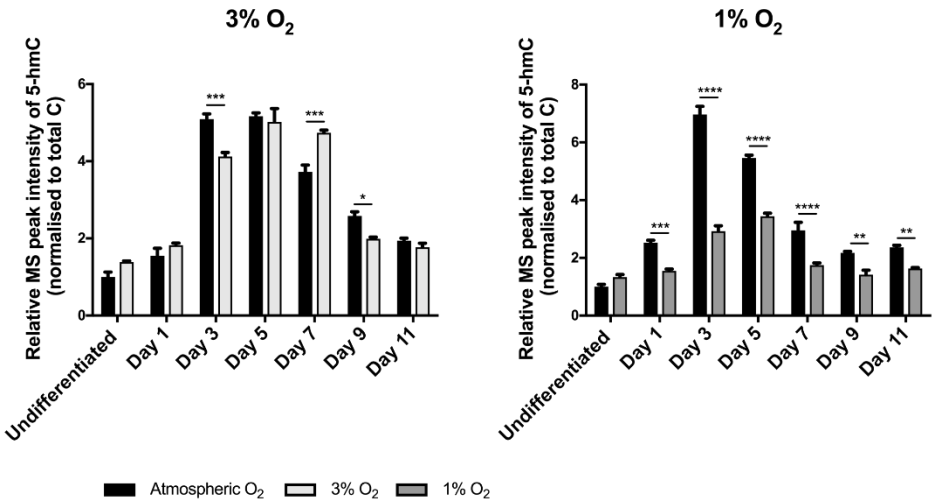


Figure 4:4 Mass-spectrometry confirms reduced hydroxylation in 1% O_2, compared to atmospheric O_2, during mESC differentiation. mESCs were differentiated over an 11 day time course in atmospheric, 3% and 1% O_2 and DNA harvested for C, 5-mC and 5-hmC quantitation by mass-spectrometry analysis. Data are expressed as the mean ± SEM relative to undifferentiated mESCs exposed to atmospheric O_2. Data are normalised to total C and analysed by a two-way ANOVA with Sidak's post-hoc analysis. * p<0.05, ** p<0.01, *** p<0.001, **** p<0.0001

4.3.3 Tet1 is predominately responsible for the burst of 5-hmC in day 3-differentiated mESCs

To determine which Tet paralogue may be catalysing the observed O_2-dependent burst of hydroxylation, the mRNA expression of each Tet enzyme over the time course of mESC differentiation was assessed (Figure 4:5). mRNA expression levels of Tet1 and Tet2 declined significantly, compared to undifferentiated levels, over the differentiation time course. In contrast, Tet3 mRNA expression increased between days 3 and 7 to give a 6.5-fold induction above undifferentiated levels, and this increase was maintained at day 11. Therefore, Tet1 and Tet2 mRNA decreased concomitant with the burst of 5-hmC at day 3, while Tet3 expression remained low until after the time point at which 5-hmC levels had peaked. Thus, neither Tet mRNA expression profiles could explain the burst of hydroxylation at day 3. Assessment of the relative expression levels of each Tet paralogue at day 3 revealed Tet1 was approximately 40-fold and 250-fold more highly expressed that Tet2 and Tet3 respectively (Figure 4:6). Hence, Tet1 being the most highly expressed suggests it is most likely to be responsible for the burst of hydroxylation at this time point.

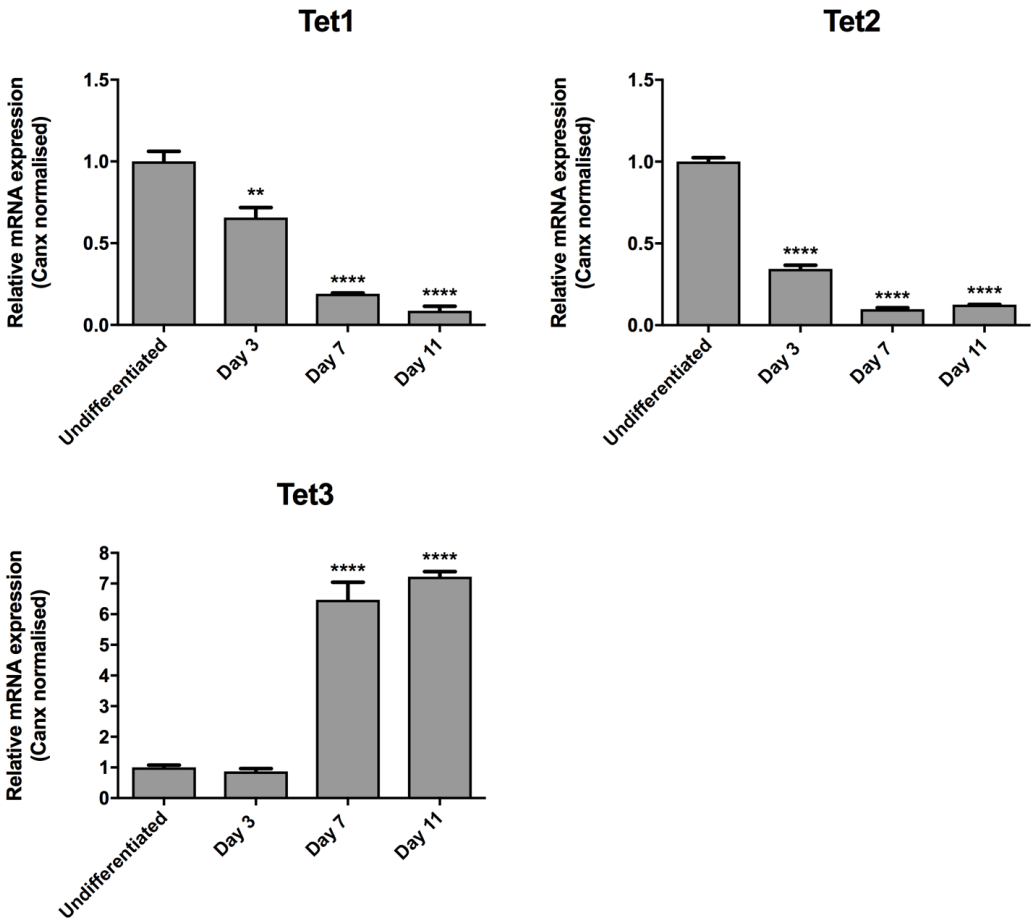


Figure 4:5 Tet1 and Tet2 expression is decreased, whereas Tet3 expression is increased during mESC differentiation. The expression of Tet1, Tet2 and Tet3 was assessed during an 11 day time course of mESC differentiation by QPCR analysis. Data are expressed as the mean ± SEM relative to undifferentiated mESCs, and analysed by a one-way ANOVA with Dunnett's post-hoc anlysis. ** $p<0.01$, **** $p<0.0001$

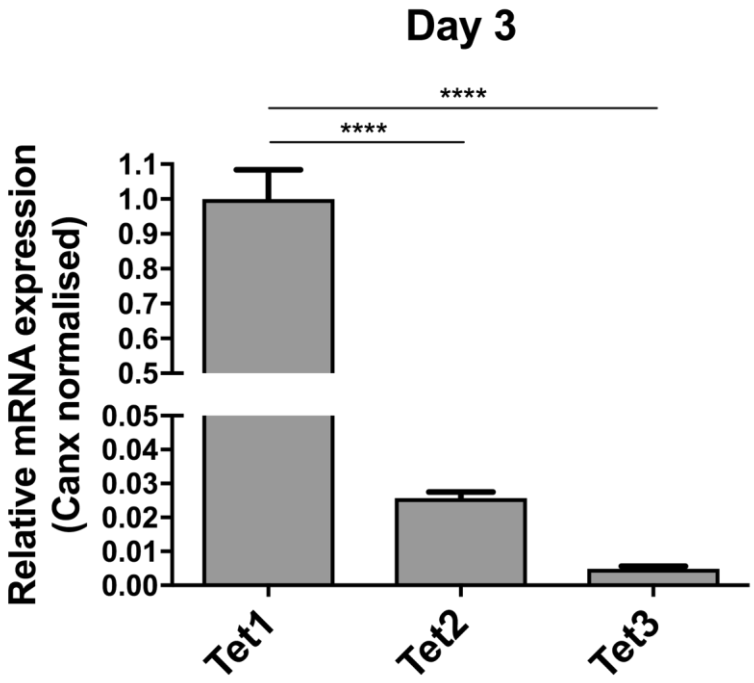


Figure 4:6 Tet1 is the most abundent isoform in day 3-differentiated mESCs. The relative expression levels of Tet1, Tet2 and Tet3 in day 3-differentiated mESCs was assessed by QPCR analysis. Data are expressed as the mean ± SEM and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. **** p<0.0001

To show functionally which Tet enzyme is responsible for mediating the burst of 5-hmC, genetically-ablated Tet mESC lines were differentiated for 3 days and levels of 5-hmC were detected by both dot-blot (Figure 4:7) and mass-spectrometry (Figure 4:8) analysis. Firstly, it was apparent that the induction of 5-hmC in shRNA control samples between undifferentiated and day 3 mESCs is consistent with the 5-hmC profile shown in Figure 4:1 and Figure 4:2. The silencing of each Tet paralogue revealed that Tet1, but not Tet2 or Tet3, reduced significantly the burst of 5-hmC at day 3, compared to shRNA control cells. This result was confirmed through mass-spectrometry, but in addition a small but significant drop in 5-hmC levels in Tet2- and Tet3- ablated mESCs compared, to shRNA controls was also identified. This was likely due to the highly accurate analytical method, resulting in little variability, leading to small errors between samples. However, it remained apparent that Tet1 ablation induced the greatest blunting effect of the burst of 5-hmC. Overall these data confirm that Tet1 is at least in part responsible for this transient burst of increased 5-hmC.

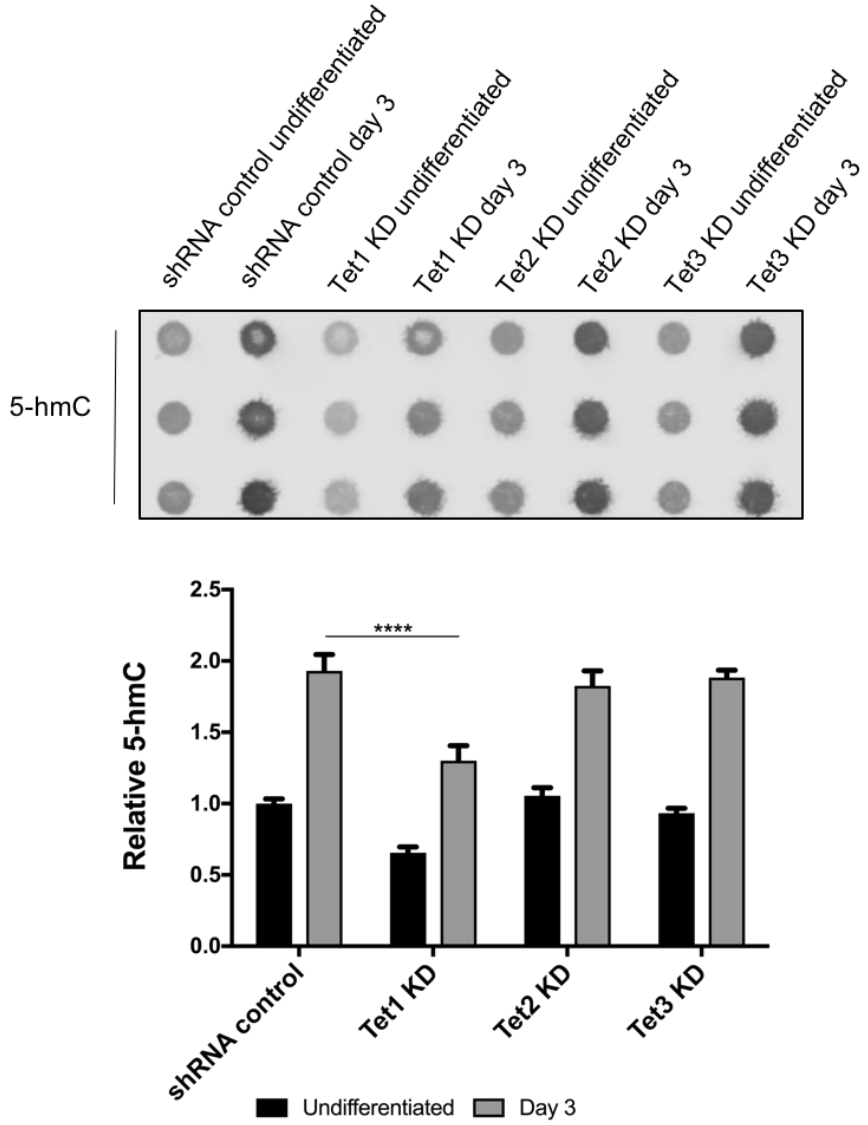


Figure 4:7 Tet1 is predominately responsible for the day 3 burst of 5-hmC in differentiating mESCs. 5-hmC levels were detected by dot-blot analysis in undifferentiated and day 3-differentiated shRNA control mESCs or mESCs silenced for either Tet1, Tet2 or Tet3. Data are expressed as the mean ± SEM relative to undifferentiated shRNA control cells, and analysed by a two-way ANOVA with Dunnett's post-hoc analysis. **** $p<0.0001$. KD: knockdown.

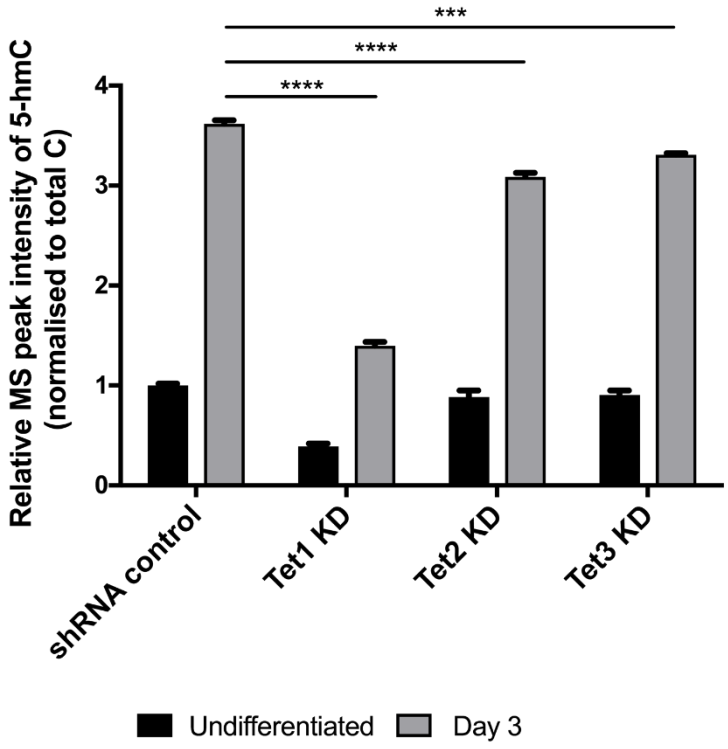


Figure 4:8 Mass-spectrometry analysis confirmed Tet1 is predominantly responsible for the hydroxylation in day 3-differentiated mESCs. C, 5-mC and 5-hmC levels were detected by mass spectrometry in undifferentiated and day 3-differentiated shRNA control mESCs or mESCs silenced for either Tet1, Tet2 or Tet3. Data are expressed as the mean ± SEM relative to undifferentiated shRNA control cells. Data are normalised to total C and analysed by a two-way ANOVA with Dunnett's post-hoc analysis. **** $p<0.001$, **** $p<0.0001$. KD: knockdown.

4.3.4 Tet1 ablation or 1% O_2 blunts the transient burst of 5-hmC to equivalent levels

As previously shown in section 3.3.6, the activity of ectopically-expressed Tet1 in HEK cells was found to be inhibited at O_2 concentrations of 3% and less. Thus, to further investigate whether Tet1 could be responsible for mediating the O_2-dependent burst of 5-hmC, levels of hydroxylation at day 3 in Tet1 KD and 1% O_2 treated mESCs were measured (Figure 4:9). These data demonstrated that the transient burst of 5-hmC is reduced to equivalent levels by depletion of Tet1 or by 1% O_2. However, combination of Tet1 ablation and 1% O_2 induced a cumulative inhibitory effect, resulting in a reduction of 5-hmC to levels to those detected in undifferentiated mESCs. 1% O_2 inhibits Tet1 activity to 50% (Figure 3:23), compared to atmospheric O_2, which could explain why 5-hmC remains elevated at day 3, above undifferentiated levels. Furthermore, Tet1 KD failed to result in the complete loss of Tet1 protein (Figure 3:28), therefore a combination of both 1% O_2 and Tet1 silencing would likely have a cumulative effect.

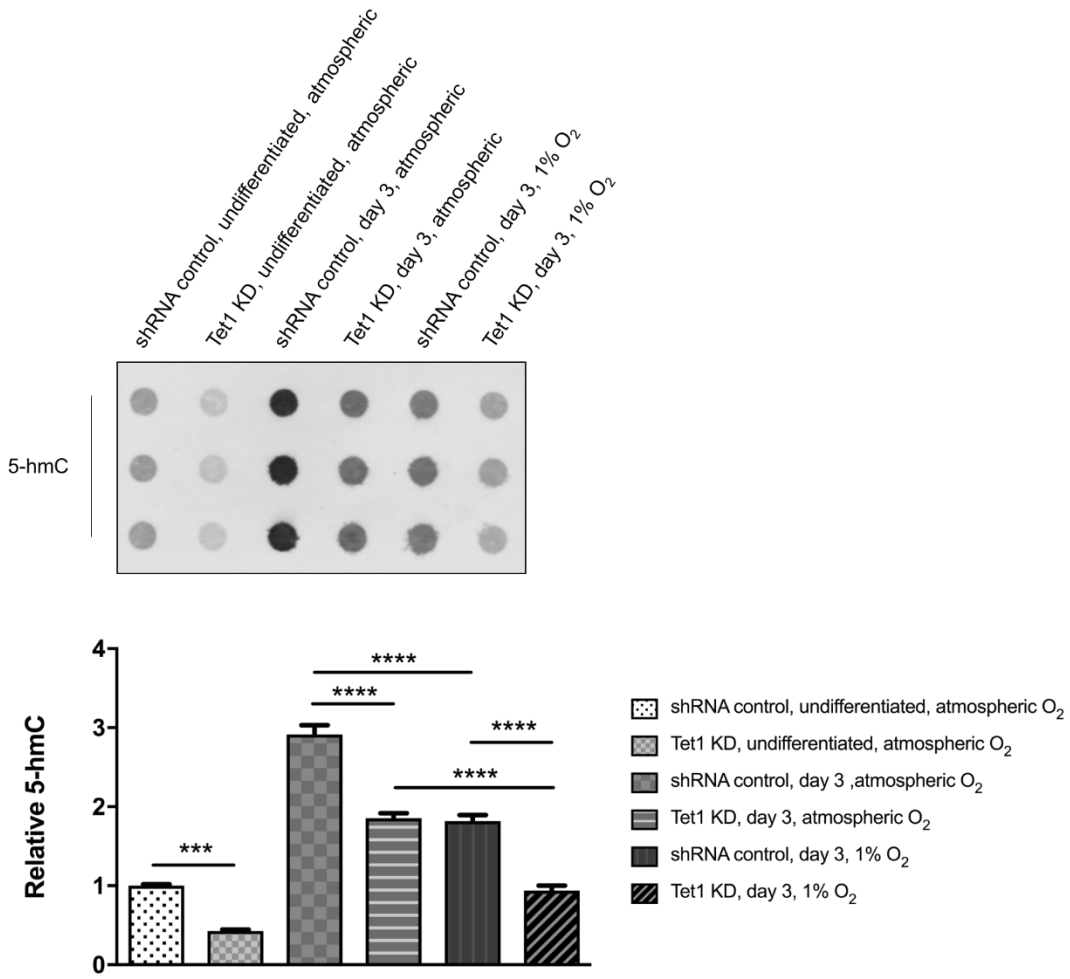


Figure 4:9 Tet1 ablation and 1% O_2 reduces 5-hmC to equivalent levels in day 3-differentiated mESCs. 5-hmC levels were quantified by dot-blot analysis for shRNA control or Tet1-depleted mESCs in either their undifferentiated state or after 3 days of differentiation in atmospheric and 1% O_2. Data are expressed as the mean ± SEM relative to shRNA control undifferentiated cells, and analysed by a one-way ANOVA with Tukey's post-hoc analysis. *** p<0.001, ***** p<0.0001. KD: knockdown.

4.3.5 Ascorbic acid increases 5-hmC in undifferentiated, but not differentiating, mESCs

The transient burst of 5-hmC did not correlate explicitly with an increase in mRNA expression of any of the Tet paralogues. Hence the activity of the Tet enzyme(s) may be increased to generate more 5-hmC at day 3. Aside from O_2 and TCA metabolites, the activity of Tets can also be upregulated by AA[125]. To confirm this, undifferentiated mESCs were exposed to increasing AA concentrations and 5-hmC levels measured (Figure 4:10). AA treatment resulted in a dose-dependent increase in 5-hmC over a specific range, reaching a maximal 2-fold induction, compared to untreated cells. Calculating the 5-hmC:5-mC ratio revealed 1 mM AA gave maximal 5-hmC elevation.

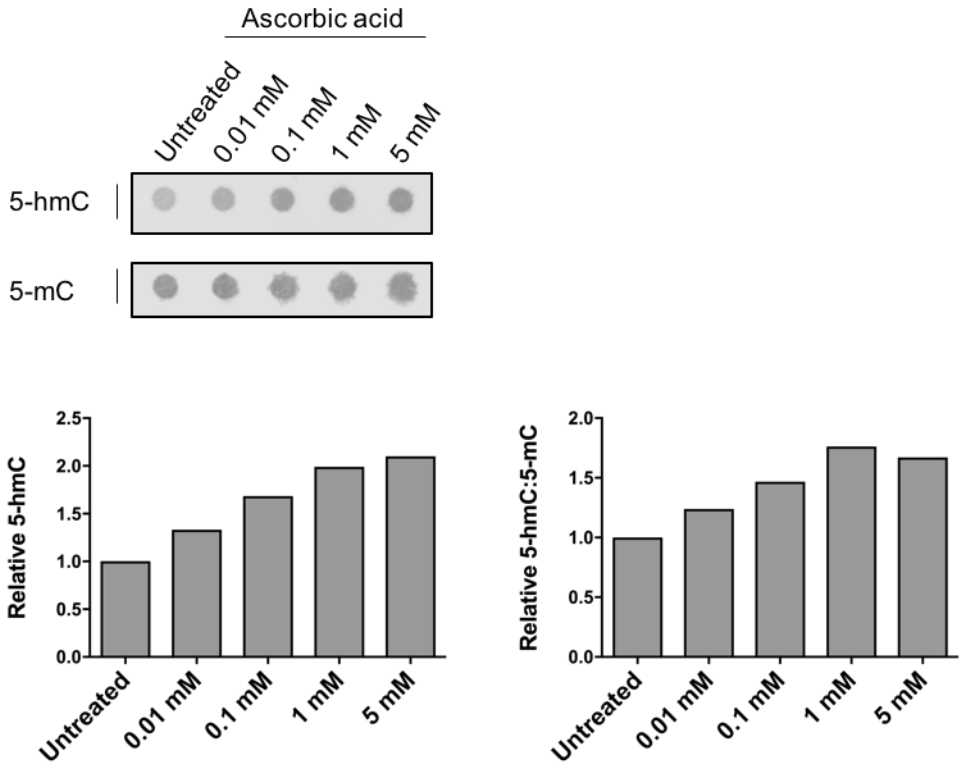


Figure 4:10 Ascorbic acid induces a dose-dependent increase in 5-hmC in undifferentiated mESCs. Undifferentiated mESCs were treated with increasing concentrations of AA (0.01 mM - 5 mM) for 24 hrs. DNA was extracted for 5-hmC quantification by dot-blot analysis.

To determine whether the transient burst of 5-hmC could be further potentiated by exogenous AA, mESCs were differentiated for 3 days in the presence of 1 mM AA (Figure 4:11). No significant increase in 5-hmC was detected between untreated and treated samples in differentiating mESCs.

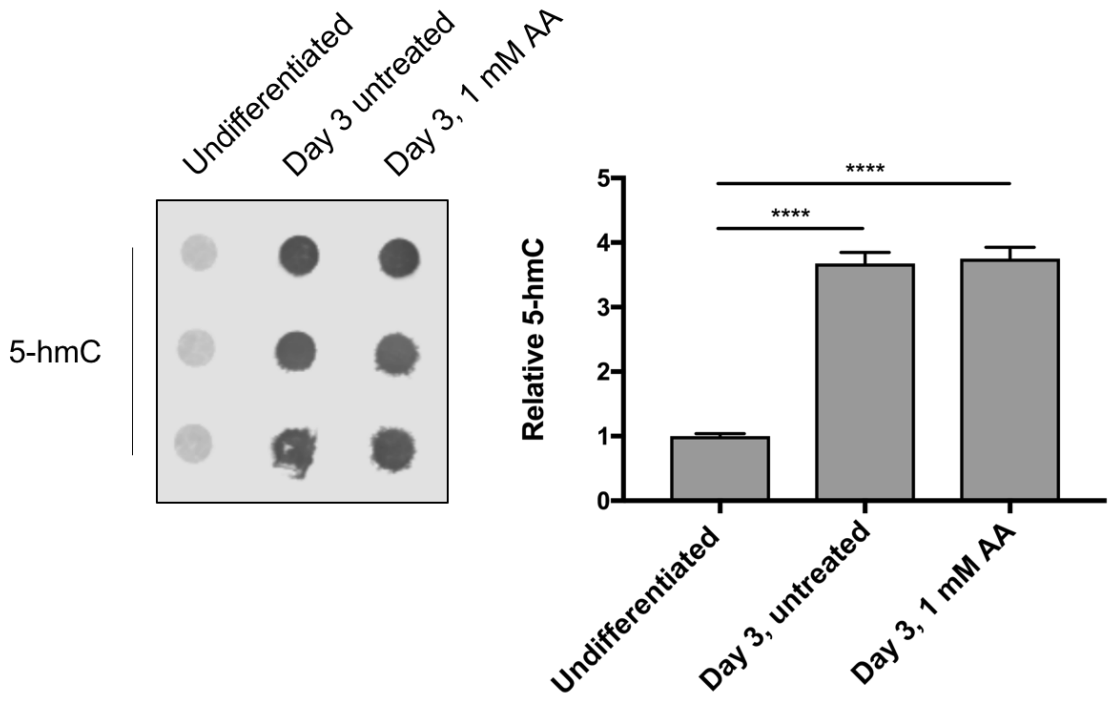


Figure 4:11 Ascorbic acid fails to potentiate the day 3 induction of 5-hmC in mESCs. mESCs were differentiated in the absence and presence of 1 mM AA for 3 days. DNA was extracted and 5-hmC levels quantified by dot-blot analysis. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by a one-way ANOVA with Tukey's post-hoc analysis. **** $p<0.0001$

4.3.6 Vitamin C transport and synthesis

Vitamin C exists in its reduced form, AA, as well as its oxidised form, DHA[467]. Transport of both forms across the membrane occurs *via* 2 distinct groups of transporter proteins. To establish whether there may be a potential influx in either form of vitamin C during differentiation, which might account for the observed transient burst of 5-hmC, the mRNA expression of each transporter across the time course of mESC differentiation was established. Svct1 and Svct2, responsible for the transport of AA[468,469], demonstrated distinct expression profiles. Svct1 showed a transient burst at day 7, compared to undifferentiated cells, whereas no expression changes for Svct2 were detected over the time course tested (Figure 4:12). Overall, it can be concluded that expression of these transporters would not account for the increased Tet activity *via* increased AA transport.

DHA transport is facilitated by the Glut transporters, specifically Glut1,3 and 4[470]. Glut1 and 3 showed similar expression profiles (Figure 4:13), with a burst of expression detected at day 7. It can be noted Glut1 was significantly increased at all differentiation time points, compared to undifferentiated cells. Glut 4 mRNA was barely detectable (data not shown). The expression profile of these transporters also does not correlate with the burst of 5-hmC at day 3. It should be recognised that in plasma AA is only detectable[471], therefore Glut transporters are not required for transport of DHA from the maternal circulation, but instead may be of importance intracellularly within the early embryo.

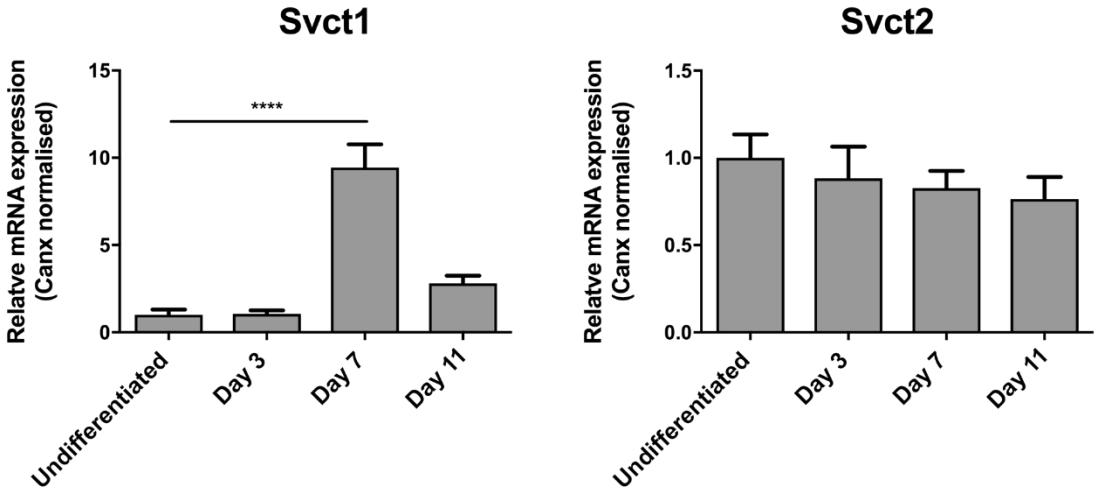


Figure 4:12 Svct1 mRNA expression shows a transient burst in day 7-differentianted mESCs. The expression of AA transporters Svct1 and Svct2 over an 11 day time course of mESCs differentiation was detected by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. **** p<0.0001

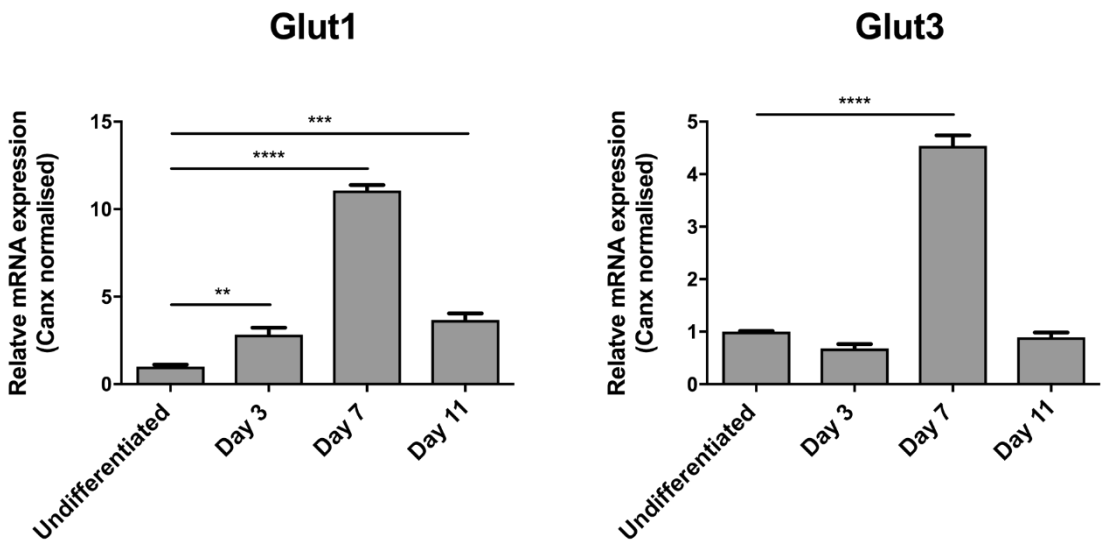


Figure 4:13 Glut1 and Glut3 mRNA expression peaks in day 7-differentiated mESCs. The expression of DHA transporters Glut 1 and Glut 3 over an 11 day time course of mESCs differentiation was detected by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

Endogenous AA synthesis occurs in mice by L-gulonolactone-γ-oxidase (Gulo), which converts L-gulonolactone, formed from D-glucose, to L-ascorbic acid[472]. Note, this enzyme is not present in man or guinea pigs and therefore AA must be acquired from the diet[473,474]. An increase in mRNA expression of Gulo may infer to a greater production of AA, which therefore may function to increase Tet activity and account for the transient 5-hmC burst in day 3-differentiated mESCs. However, Gulo expression was only shown to increase significantly from undifferentiated mESCs at day 11 (Figure 4:14), suggesting this hypothesis is unlikely.

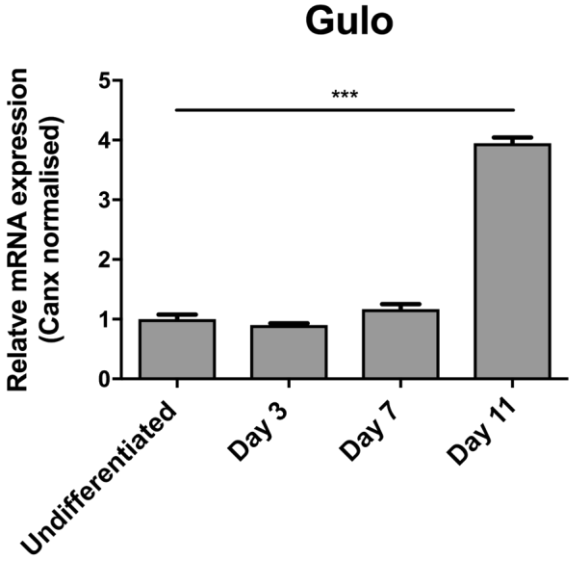


Figure 4:14 L-gulonolactone-γ-oxidase mRNA expression is induced in day 11-differentiated mESCs. The mRNA expression of the AA synthesis gene Gulo was assessed over an 11-day time course of mESC differentiation by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. *** p<0.001

4.3.7 Manipulation of 5-hmC levels by a reduction in AA levels induced by buthionine sulfoximine and phloretin treatment

To investigate further whether 5-hmC levels might be dependent upon the regulation of Tet activity by AA, the effect of inhibition of AA uptake and DHA reduction by phloretin[475] and BSO[476] respectively were studied. Phloretin prevents AA uptake through inhibition of the sodium-dependent transporters, Svct1 and Svct2[476] (and also Glut2[477]). In comparison, BSO inhibits γ-glutamylcysteine synthetase, which is responsible for GSH synthesis[478]. GSH reduces and recycles DHA to AA[479], and thus makes AA available as a cofactor for Tet enzymes (and other 2-OG-dependent dioxygenases).

Undifferentiated mESCs were treated with 1 mM BSO or 200 μM phloretin and separately pre-treated with either BSO or phloretin for 3 hrs before the addition of 1 mM AA (Figure 4:15). Cells were harvested 24 hrs after treatments. BSO treatment had no significant effect on 5-hmC levels compared to untreated cells and perhaps as expected, had no inhibitory effects in combination with exogenous AA treatment. In comparison, phloretin also had no significant effect on 5-hmC levels, compared to its DMSO vehicle control. However, in combination with AA, phloretin did blunt the increased 5-hmC.

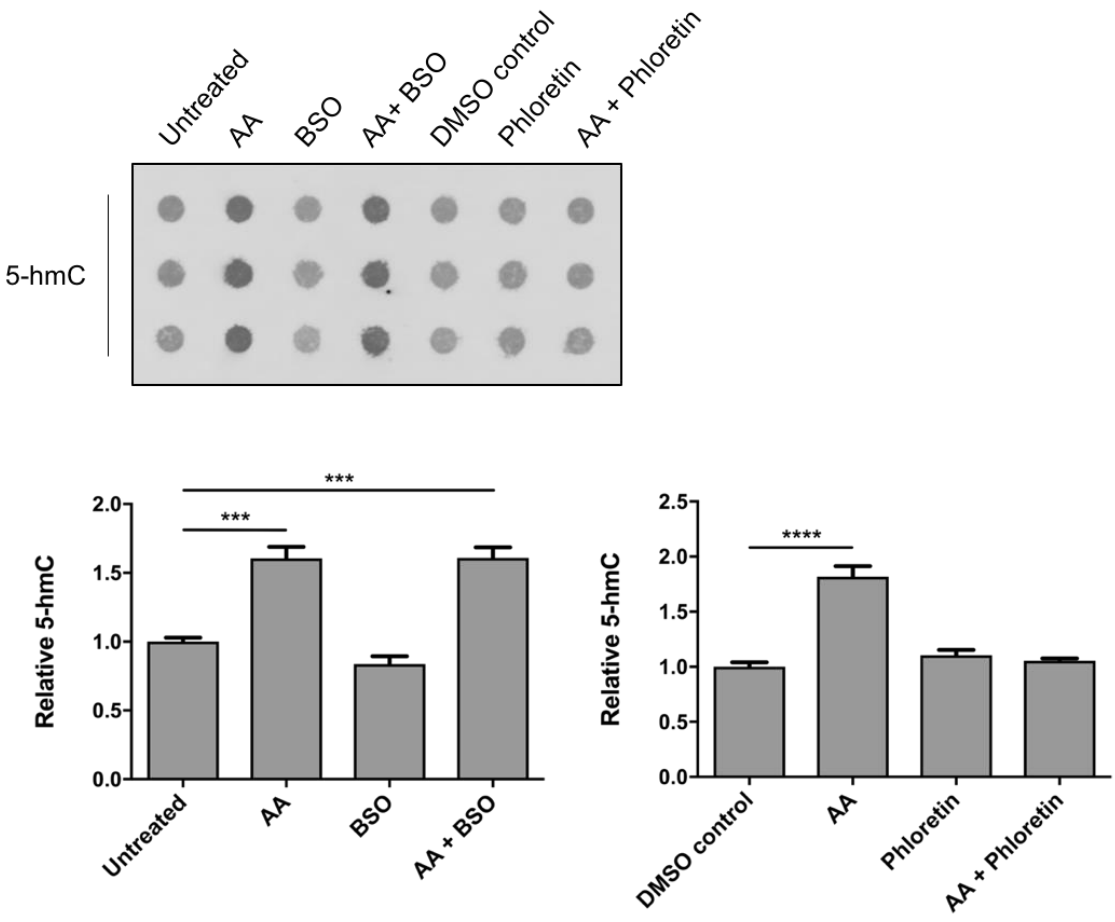


Figure 4:15 Phloretin inhibits uptake of Ascorbic Acid in undifferentiated mESCs. Undifferentiated mESCs were treated with 1 mM AA, or left as an untreated control, for 3 hrs, prior to the addition of 1 mM BSO or 200 μM phloretin. DNA was extracted 24 hrs post AA addition and 5-hmC levels quantified by dot-blot analysis. Data are expressed as the mean ± SEM relative to untreated or DMSO vehicle control, for BSO and phloretin treatments respectively. Data are analysed by a one-way ANOVA with Dunnett's post-hoc analysis. *** $p<0.001$, **** $p<0.0001$

To determine the effect of AA upon the transient burst of 5-hmC, mESCs were differentiated for 3 days in the presence of BSO or phloretin (Figure 4:16). BSO treatment demonstrated a modest dose-dependent inhibitory effect on 5-hmC levels, compared to untreated cells. Despite a minimal effect, these data suggest AA may have some contribution to the burst of 5-hmC at day 3. Phloretin exerted no repressive effect on 5-hmC levels, compared to DMSO vehicle control treated cells, thereby suggesting that uptake of AA from cell medium has no effect upon 5-hmC levels at day 3.

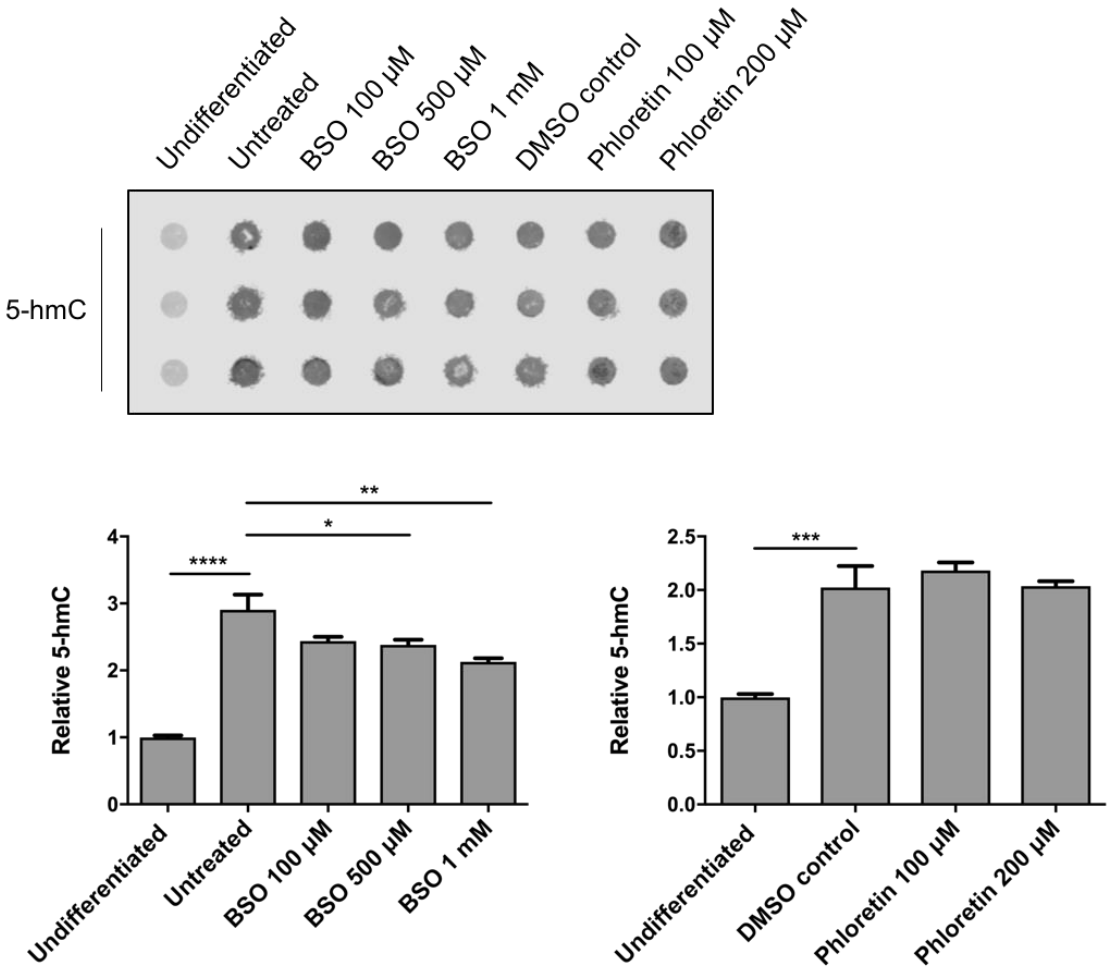


Figure 4:16 Buthionine sulfoxime modestly reduced 5-hmC levels in day 3-differentiated mESCs. mESCs were induced to differentiate in the presence of increasing concentrations of BSO (100 μM-1 mM) or phloretin (100 μM-200μM) for 3 days. DNA was extracted and 5-hmC levels quantified by dot-blot analysis. Data are expressed as the mean ± SEM relative to undifferentiated mESCs. Data are analysed by a one-way ANOVA with Dunnett's post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

4.3.8 Differing AA content of mESC maintenance and differentiating medium had no influence upon the burst of 5-hmC in day 3-differentiated cells

mESCs were maintained and differentiated in 2 distinct types of medium (Chapter 2: Material and Methods). While the levels of AA in maintenance medium were low (8.6 μM in Advance DMEMF-12), the amount contained within the KO serum replacement, used for differentiation experiments is understood to be considerably higher (Invitrogen pers. comm.), although the exact amounts are proprietary[480]. To ensure that the burst of 5-hmC in day 3-differentiated mESCs was not due to changes in cell medium and thus potentially AA concentration, undifferentiated mESCs were cultured in maintenance medium or maintenance medium supplemented with 15% KO serum replacement (as used in differentiation studies) instead of EmbryoMax® FBS. In addition, cells maintained under these 2 serum-differing conditions were induced to differentiate to day 3, under normal differentiating KO serum conditions. In this way, the effect of switching from EmbryoMax® FBS to KO serum (and hence changing AA levels) upon 5-hmC levels was determined.

KO serum elevated 5-hmC, compared to EmbryoMax®, in undifferentiated mESCs, but not to the same extent as 1 mM AA treatment (Figure 4:17). At day 3 of mESC differentiation the burst of hydroxylation was evident for cells maintained undifferentiated in either serum condition, compared to undifferentiated levels (Figure 4:17). Crucially no significant difference was observed in the relative 5-hmC levels between these serum conditions. Therefore, it can be concluded that differing AA amounts in cell medium has no influence upon the burst of 5-hmC at day 3. This is also supported by the data from the AA uptake inhibitor phloretin, shown in Figure 4:16.

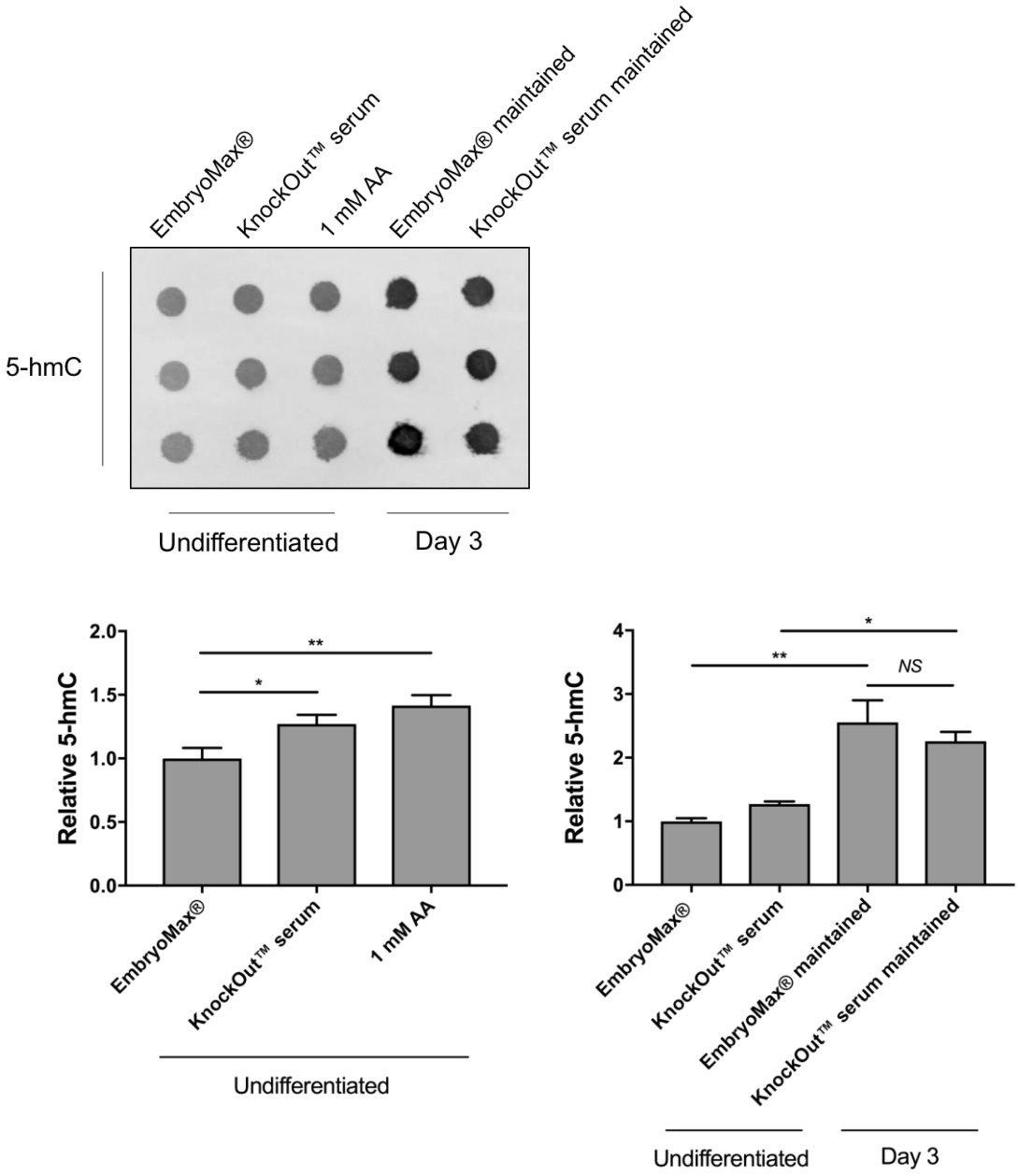


Figure 4:17 KnockOut™ serum moderately increases 5-hmC in undifferentiated mESCs, but has no influence on the burst of hydroxylation in day 3-differentiated mESCs. 5-hmC levels were quantified from undifferentiated mESCs cultured in maintenance medium (containing EmbryoMax® FBS) or maintenance medium supplemented with 15% KO serum instead of EmbryoMax®, and compared to 1 mM AA treated mESCs (maintained under normal EmbryoMax® conditions). In addition, 5-hmC levels were determined for mESCs maintained under these 2 serum conditions after 3 days of differentiation (under normal KO serum differentiating conditions). Dot-blot quantified data was expressed as the mean ± SEM relative to normally maintained (EmbryoMax®) undifferentiated cells. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. * $p<0.05$, ** $p<0.01$

4.3.9 The transient day 3 burst of 5-hmC is predominately targeted to a Tet3 promoter region

To identify the genomic targets of the burst of 5-hmC at day 3 of mESC differentiation, a hMeDIP-seq study was conducted (workflow schematically shown in (Figure 4:18)). 610 gene loci were identified to be enriched greater than 2-fold compared to total genomic input DNA (see supplementary data table in Burr, S *et al* (2017)[481]). Unexpectedly the largest enrichment, which was more than 7-fold higher than other identified loci, was mapped to a CG-rich promoter region of the Tet3 gene, which also comprised the B230319CO9Rik sequence tag (visualised in Figure 4:19). Functional annotation clustering of the gene list using the DAVID Gene Functional Annotation Tool (https://david.ncifcrf.gov) revealed the most enriched cluster to contain 55 genes that were associated with developmental and differentiation pathways (see screen shot in Figure 4:20A). In addition, some clusters highlighted a selection of genes involved in the development of specific tissues, including cardiac morphogenesis (see screen shot in Figure 4:20B) and neural development (see screen shot in Figure 4:20C).

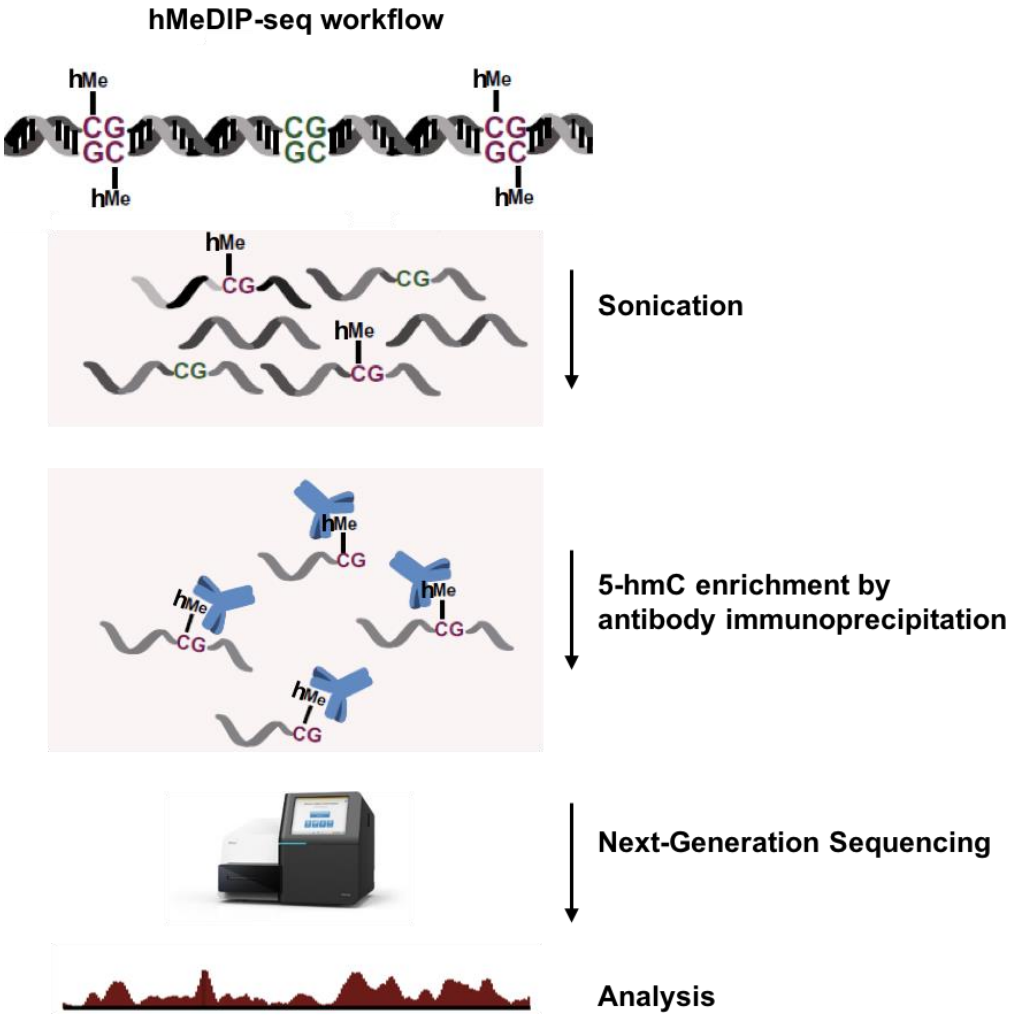


Figure 4:18 Workflow for 5-hmC-DNA immunoprecipitation-sequencing. Isolated genomic DNA was sonicated into 200-600 bp fragments and 5-hmC DNA was captured and purified by antibody immunoprecipitation. Next-generation sequencing libraries were prepared from 5-hmC enriched DNA (this involved the following steps: dA tailing, adaptor ligation, size selection, barcoding and PCR amplification). The library pools were sequenced and analysed to identify the genomic locations containg the 5-hmC mark. Adapted from Li, D *et al*. Methods. 2015; 72:29-40[588].

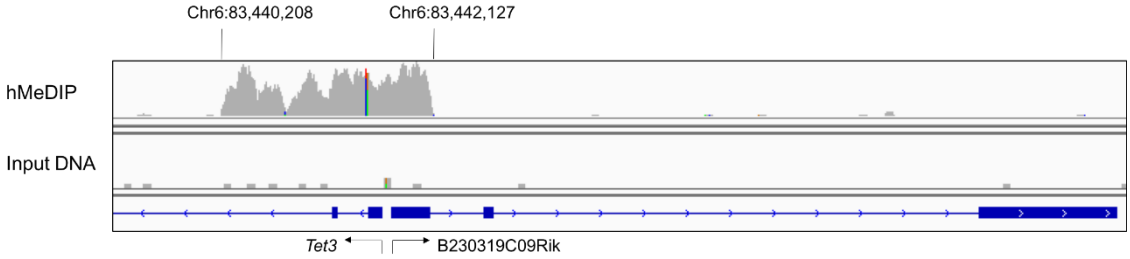


Figure 4:19 5-hmC-DNA immunprecipitation-sequencing revealed Tet3 is enriched for hydroxylation in day 3-differentiated mESCs. hMeDIP-seq conducted on day 3-differentiated mESCs revealed Tet3 and also the non-coding B230319C09Rik transcript was a target for hydroxylation. Data are visualised by an annotated screen-shot from the Integrative Genomics Viewer.

A

Annotation Cluster 1	Enrichment Score: 3.17			Count	P_Value	Benjamini
UP_KEYWORDS	Developmental protein	RT		44	2.6E-5	2.0E-3
GOTERM_BP_DIRECT	multicellular organism development	RT		45	2.4E-4	2.6E-1
UP_KEYWORDS	Differentiation	RT		27	3.4E-3	6.7E-2
GOTERM_BP_DIRECT	cell differentiation	RT		31	1.0E-2	6.3E-1

B

Annotation Cluster 6	Enrichment Score: 1.99			Count	P_Value	Benjamini
GOTERM_BP_DIRECT	atrioventricular valve morphogenesis	RT		5	3.9E-4	2.8E-1
GOTERM_BP_DIRECT	positive regulation of epithelial to mesenchymal transition	RT		5	9.2E-3	6.5E-1
GOTERM_BP_DIRECT	cardiac left ventricle morphogenesis	RT		3	4.5E-2	8.7E-1
GOTERM_BP_DIRECT	ventricular septum morphogenesis	RT		4	6.6E-2	9.0E-1

Annotation Cluster 8	Enrichment Score: 1.93			Count	P_Value	Benjamini
GOTERM_BP_DIRECT	atrioventricular valve morphogenesis	RT		5	3.9E-4	2.8E-1
GOTERM_BP_DIRECT	endocardial cushion development	RT		3	3.9E-2	8.6E-1
GOTERM_BP_DIRECT	positive regulation of cardiac muscle cell proliferation	RT		3	1.1E-1	9.4E-1

Annotation Cluster 10	Enrichment Score: 1.85			Count	P_Value	Benjamini
GOTERM_BP_DIRECT	atrioventricular valve morphogenesis	RT		5	3.9E-4	2.8E-1
GOTERM_BP_DIRECT	atrial septum morphogenesis	RT		4	4.3E-3	5.4E-1
GOTERM_BP_DIRECT	positive regulation of cardioblast differentiation	RT		3	5.7E-3	6.2E-1
GOTERM_BP_DIRECT	response to estrogen	RT		7	1.0E-2	6.7E-1
GOTERM_BP_DIRECT	cardiac right ventricle morphogenesis	RT		3	3.9E-2	8.6E-1
GOTERM_BP_DIRECT	outflow tract morphogenesis	RT		4	1.5E-1	9.6E-1
GOTERM_BP_DIRECT	heart morphogenesis	RT		4	2.1E-1	9.8E-1

C

Annotation Cluster 2	Enrichment Score: 2.79			Count	P_Value	Benjamini
GOTERM_CC_DIRECT	cell junction	RT		34	2.2E-4	2.9E-2
UP_KEYWORDS	Cell junction	RT		31	2.7E-4	1.0E-2
GOTERM_CC_DIRECT	synapse	RT		25	1.0E-3	7.7E-2
UP_KEYWORDS	Postsynaptic cell membrane	RT		12	2.3E-3	5.2E-2
GOTERM_CC_DIRECT	postsynaptic membrane	RT		14	2.5E-3	1.5E-1
UP_KEYWORDS	Synapse	RT		18	3.4E-3	7.1E-2
GOTERM_CC_DIRECT	postsynaptic density	RT		12	2.7E-2	4.2E-1

Annotation Cluster 26	Enrichment Score: 1.15			Count	P_Value	Benjamini
GOTERM_BP_DIRECT	positive regulation of dendritic spine development	RT		4	3.3E-2	8.4E-1
GOTERM_CC_DIRECT	dendritic spine	RT		8	6.2E-2	6.0E-1
GOTERM_CC_DIRECT	dendritic shaft	RT		4	1.7E-1	7.9E-1

Figure 4:20 Functional annotation clustering of 601 gene loci identified from 5-hmC-DNA immunprecipitation-sequencing in day 3-differentiated mESCs. 601 gene loci were identified from hMeDIP-seq that displayed a greater than 2-fold enrichment compared to input DNA. Functional annotation clustering of these loci using the DAVID Gene Functional Annotation Tool (accessed at https://david.ncifcrf.gov) revealed (A) developmental associated genes, comprising 55 loci, was the most enriched cluster. Other annotated clusters included tissue specific genes involved in (B) cardiac and (C) neural development.

4.3.10 The CG-rich Tet3 promoter region gains de-novo 5-hmC at the onset of differentiation

To establish whether this Tet3 promoter region acquires 5-hmC as a result of differentiation, or alternatively is associated with this mark in undifferentiated mESCs, genomic DNA from undifferentiated and day 3-differentiated mESCs was subjected to bisulfite and oxidative-bisulfite treatment (TrueMethyl® Whole Genome kit). A 200 bp fragment containing 3 CpGs within the same Tet3 promoter region as that enriched for 5-hmC, shown in Figure 4:19, was amplified by PCR and the amplicon sequenced to assesses the methylation and hydroxymethylation status (Figure 4:21). Bisulfite treatment acts to convert all unmethylated C residues to uracil (U), which are subsequently converted to thymine (T) in the PCR reaction. Methylated and hydroxymethylated residues are protected from conversion to U and thus remain as a C. In undifferentiated mESCs all 3 CpG dinucleotides were devoid of methylation (and hydroxymethylation). However, in day 3-differentiated samples all CpGs were converted to C, depicted by the arrow, indicative of containing the 5-mC or 5-hmC marks. To distinguish between these modifications an oxidative bisulfite step was implemented. This oxidative step converts C to U and oxidises 5-hmC to 5-fC, but has no effect on 5-mC. After the PCR reaction 5-fC (and U) is converted to T while 5-mC remains as a C residue. 2/3 CpGs within this PCR amplicon were converted to T, indicating the presence of the 5-hmC mark. The remaining CpG, as represented by the arrow, persisted as a C residue, suggesting the presence of the 5-mC mark. Overall, these data show that no methylation is present within this Tet3 promoter region in undifferentiated mESC, but once induced to differentiate, by day 3, this region showed the presence of 5-hmC (and 5-mC).

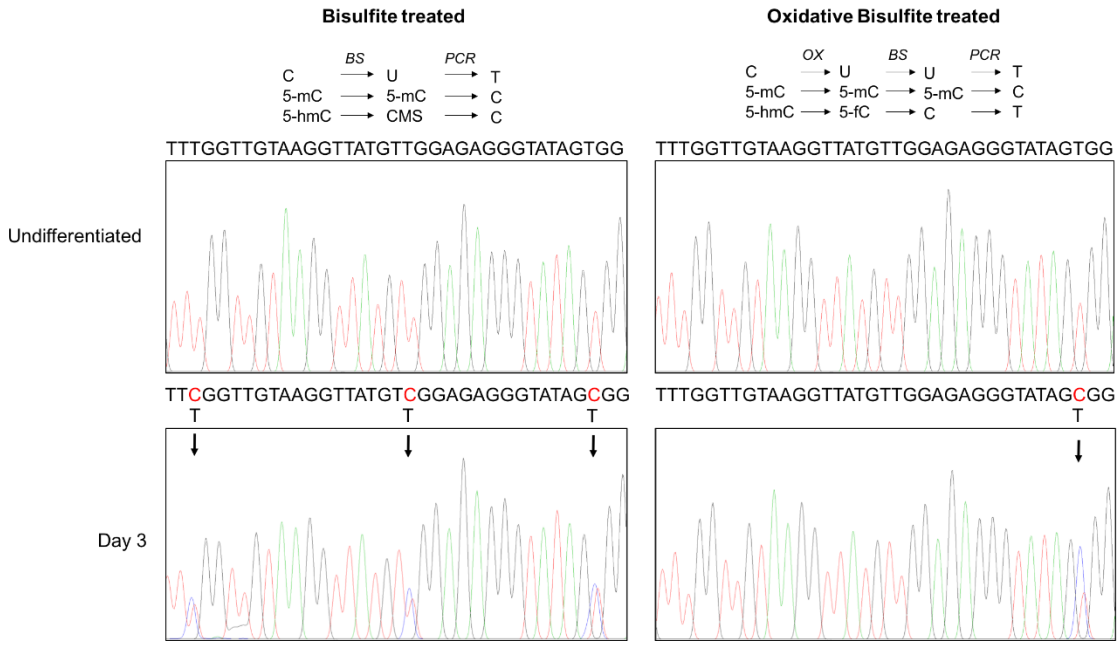


Figure 4:21 Bisulfite and Oxidative Bisulfite treatment revealed enrichment of hydroxylation within the Tet3 promoter region from undifferentiated to day 3-differentiated mESCs. DNA extracted from undifferentiated and day 3-differentiated mESCs were subjected to bisulfite and oxidative bisulfite treatment (TrueMethyl® Whole Genome kit). A 200 bp fragment containing 3 CpG dinucleotides within a Tet3 promoter region was amplified from these samples and its methylation status interrogated by Sanger sequencing. Shown here are screen shots from Sanger sequencing chromatograms (4Peaks software), annotated with arrows to depict methylated CpGs (5-mC or 5-hmC) in bisulfite treated samples and the 5-mC-containing CpG in oxidative-bisulfite samples. CMS=5-methylenesulfonate.

4.3.11 Tet1 is bound to this Tet3 promoter region in undifferentiated and day 3-differentiated mESCs

Tet1 has previously been shown to mediate, at least in part, the transient burst of 5-hmC apparent within day 3-differentiated mESCs (see section 4.3.3). The most significant target of this hydroxylation was found to be a CG-rich Tet3 promoter region (see Figure 4:19), which therefore may suggest potential for an interaction between Tet1 and Tet3. To determine whether Tet1 was localised to this Tet3 promoter region a ChIP assay was performed on undifferentiated and day 3-differentiated mESCs. It was confirmed that Tet1 was bound to this Tet3 promoter region in both undifferentiated and in day 3-differentiated mESCs by semi-quantitative PCR (Figure 4:22). QPCR analysis further revealed that Tet1 was bound with equivalent occupancy to this Tet3 promoter region in undifferentiated and day 3-differentiated mESCs (Figure 4:23).

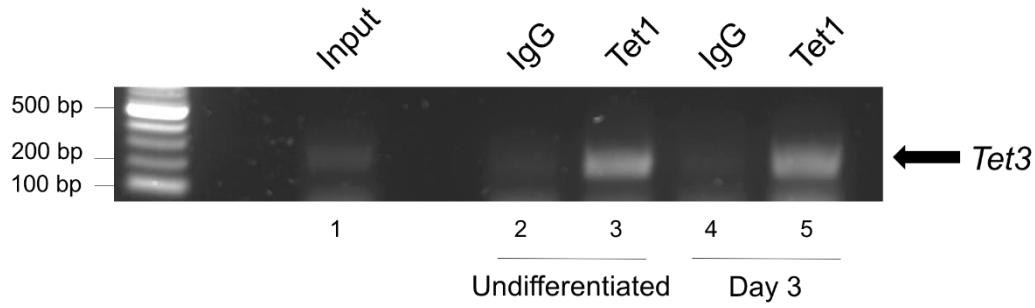


Figure 4:22 Tet1 is bound to the Tet3 promoter in undifferentiated and day 3-differentiated mESCs. Tet1 ChIP was performed on undifferentiated and day 3-differentiated mESCs samples. A fragment of the Tet3 promoter region was amplifed by PCR from input DNA, IgG control and ChIP samples. PCR products were visualised by agarose-gel electrophoresis.

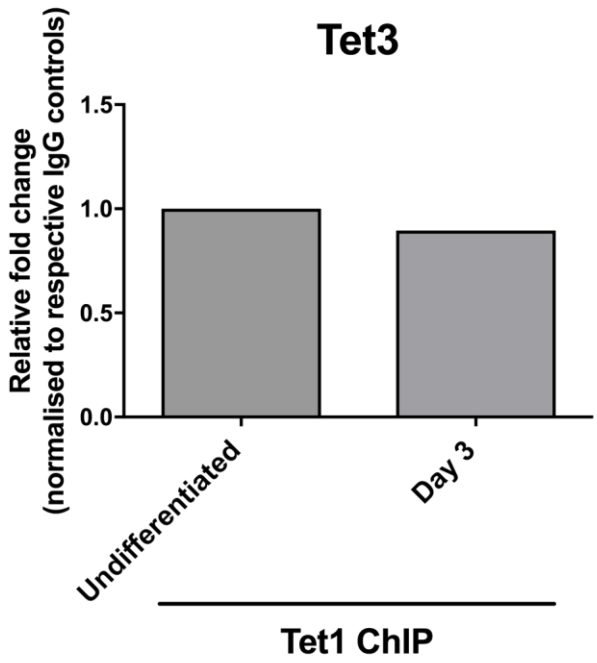


Figure 4:23 Tet1 is bound to Tet3 with equivalent occupancy in undifferentiated and day 3-differentiated mESCs. Assessment of Tet3 mRNA expression was determined by QPCR from chromatin isolated from undifferentiated and day 3-differentiated mESC samples, which had been immunoprecipitated using a Tet1-specific antibody. Data are expressed relative to undifferentiated mESCs.

4.3.12 The histone marks H3K4me3 and H3K27me3 share occupancy with Tet1 at this Tet3 promoter region

Tet1 is known to be associated with bivalent gene promoters, whereby the chromatin is characterised by the presence of both activating (H3K4me3) and repressive (H3K27me3) histone marks[194,193]. Owing to the finding that Tet1 is localised to a CG-rich Tet3 promoter region in both undifferentiated and day 3-differentiated mESCs (see section 4.3.11), the potential association with histone marks at this locus was also assessed by ChIP. It was revealed that this Tet3 promoter region was occupied by both these histone marks in undifferentiated and day 3-differentiated mESCs (Figure 4:24). At day 3, the H3K4me3 mark persisted at a similar level to that in undifferentiated mESCs (demonstrated by the equivalently intense band). However, the repressive H3K27me3 mark appeared to be decreased at day 3, compared to undifferentiated mESCs. QPCR analysis confirmed the decrease in H3K27me3 levels from undifferentiated to day 3-differentiated mESCs (Figure 4:25). However, in contrast to the results of semi-quantitative PCR analysis, H3K4me3 levels appeared increased at day 3 when assessed by QPCR.

These data identified Tet3 as a bivalent promoter, and hence the decrease in H3K27me3 (and maintenance/increase of H3K4me3) from undifferentiated to day 3-differentiated mESCs correlated with increased 5-hmC (see Figure 4:21) and the induction of Tet3 mRNA expression (see Figure 4:5).

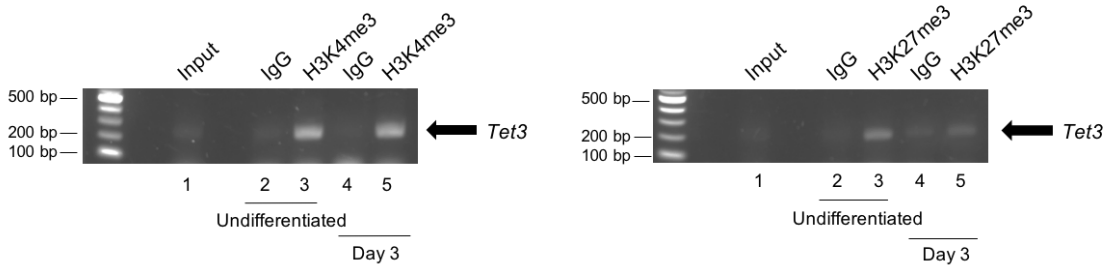


Figure 4:24 The H3K4me3 and H3K27me3 histone marks are associated with the Tet3 promoter and appear maintained or lost respectively in day 3-differentiated mESCs. H3K4me3 and H3K27me3 ChIP was performed on undifferentiated and day 3-differentiated mESCs. A fragment of the Tet3 promoter region was amplifed by PCR from input DNA, IgG control and ChIP samples. PCR products were visualised by agarose-gel electrophoresis.

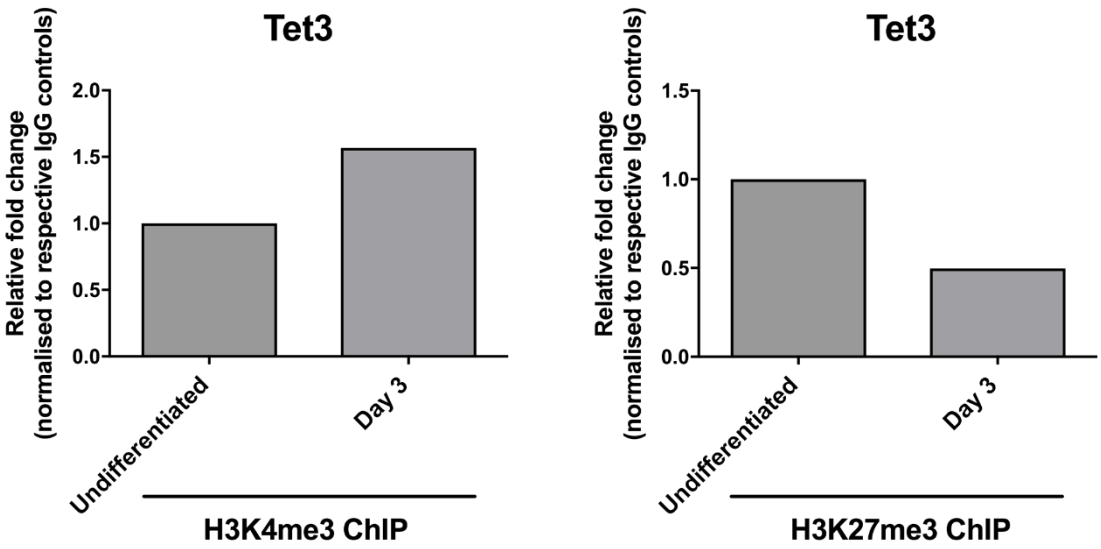


Figure 4:25 The Tet3 promoter gains H3K4me3 and loses H3K27me3 between undifferentiated and day 3-differentiated mESCs. Assessment of Tet3 mRNA expression was determined by QPCR from chromatin isolated from undifferentiated and day 3-differentiated mESC samples, which had been immunoprecipitated using H3K4me3 and H3K27me3-specific antibodies. Data are expressed relative to undifferentiated mESCs.

4.3.13 Ablation of Tet1 induced protein expression of H3K4me3 and H3K27me2

Confirmation that Tet1 is associated with activating and repressive histone marks (see section 4.3.12) led to further investigation into the potential cross-talk between these epigenetic modifications. Global changes in the protein expression of the repressive (H3K27me2/3) and activating (H3K4me3) histone marks were assessed in Tet-ablated cells in day 3-differentiated mESCs (Figure 4:26). Tet1 silencing specifically induced expression of H3K27me2 and H3K4me3. Tet1- and Tet3- ablation both trended to increase H3K27me3, but failed to reach significance. Overall this confirms the presence of complex epigenetic signalling networks, which may function to precisely control gene expression.

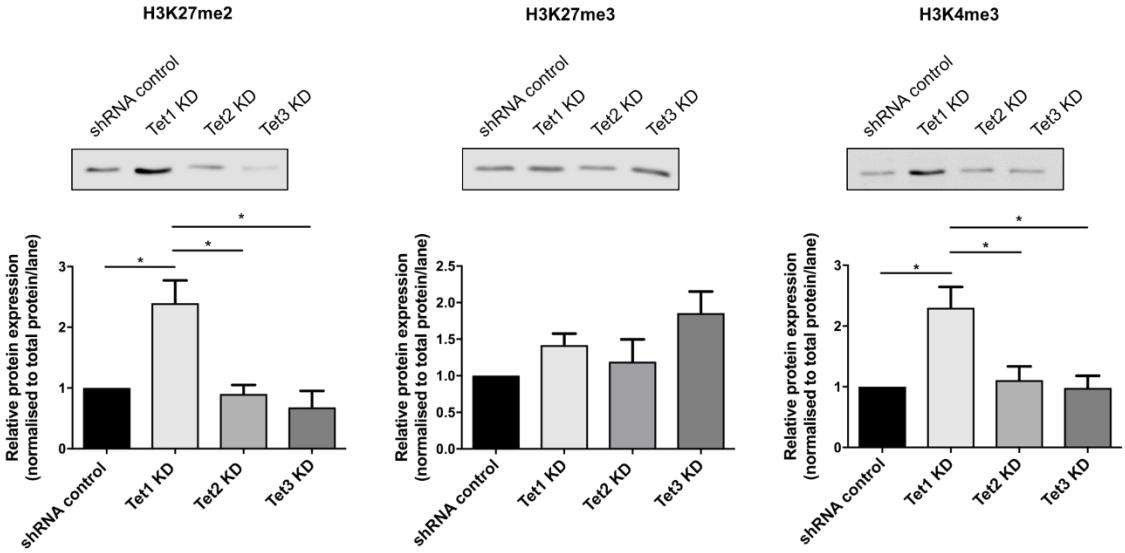


Figure 4:26 Tet1 silencing induces H3K27me2 and H3K4me3 protein expression in day 3-differentiated mESCs. shRNA control and Tet-silenced mESCs were induced to differentiate for 3 days. Protein was extracted and H3K27me2/3 and H3K4me3 expression levels were determined by Western blot analysis. Data are shown as the mean ± SEM from 3 independent experiments, and expressed relative to the shRNA control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. *p<0.05. KD: knockdown.

4.3.14 The 5-hmC mark found to be enriched within the CG-rich Tet3 promoter region is lost by differentiating mESCs in 1% O_2

Differentiation of mESCs in 1% O_2, compared to atmospheric O_2, reduced 5-hmC levels (see section 4.3.2). To determine whether the hydroxylation enriched at the Tet3 promoter region (identified in Figure 4:19) is reduced in a low [O_2], hMeDIP-PCR analysis was conducted on day 3-differentiated mESCs exposed to atmospheric and 1% O_2. Primers against the Tet3 promoter were designed to amplify close to the regions used for PCR amplification in bisulfite/oxidative-bisulfite and ChIP experiments, which encompassed the 5-hmC enriched area identified from the hMeDIP-seq study.

In atmospheric O_2, 5-hmC was as expected enriched within this Tet3 promoter region. However, this enrichment was lost in mESCs cultured in 1% O_2 (Figure 4:27). The exact implication of this with respect to Tet3 mRNA expression is investigated below (see section 4.3.15).

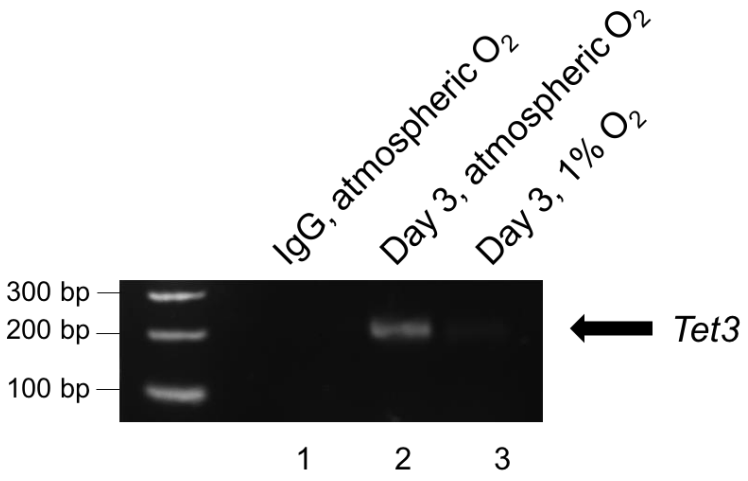


Figure 4:27 5-hmC enrichment of the Tet3 promoter is lost in 1% O_2. mESCs were differentiated for 3 days in atmospheric or 1% O_2 and DNA extracted for 5-hMeDIP-PCR. A Tet3 promoter fragment, which encompassed the area of 5-hmC enrichment identifed from the hMeDIP-seq study was amplifed by PCR. PCR products were visualised by agarose-gel electrophoresis.

4.3.15 Tet3 mRNA expression is reduced in differentiating mESCs exposed to low [O_2]

To determine whether the loss of the 5-hmC mark in 1% O_2 (Figure 4:22) impacted upon Tet3 expression, mESCs were induced to differentiate over 11 days in atmospheric, 3% or 1% O_2 and Tet3 mRNA levels assessed (Figure 4:28).

In undifferentiated mESCs, culture under low [O_2] trended to decrease Tet3 mRNA expression. However, by day 3 mRNA levels were increased under low O_2 exposed conditions. By days 7 and 11 Tet3 expression was significantly decreased in both 3 and 1% O_2, compared to atmospheric O_2. Taken together, these data propose that O_2-dependent hydroxylation of the Tet3 promoter, at day 3 of mESC differentiation, may act as a 'mark' for increased expression at a later stage of cellular differentiation. However, [O_2] is likely to affect many cellular processes, due to its intrinsic role as a cofactor for 2-OGDO containing enzymes, and thus may have additional transcriptional regulatory roles. For example, it may be hypothesised from the data shown here that at the day 3-time point [O_2] can alter the histone methylation marks (see section 4.3.16) to influence mRNA expression.

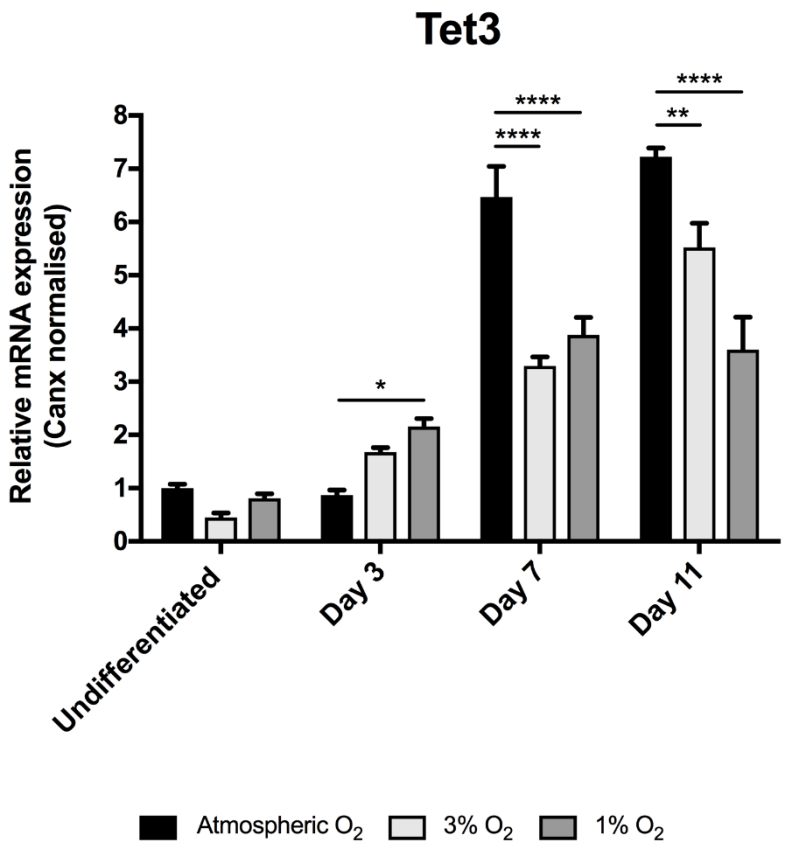


Figure 4:28 Tet3 mRNA expression is reduced in 3 and 1% O_2 at day 7- and 11-differentiated mESCs. The mRNA expression of Tet3 was assessed by QPCR in mESCs over an 11 day differentiation time course in atmospheric, 3% and 1% O_2. Data are expressed as the mean ± SEM relative to undifferentiated mESCs exposed to atmospheric O_2. Date are analysed by a two-way ANOVA with Tukey's post-hoc analysis. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$

4.3.16 1% O_2 induced protein expression of H3K27me2/3 and H3K4me3 in day 3-differentiated mESCs

Literature evidence describes a correlation of 5-hmC with the H3K27me3[482] and H3K4me3[483] histone marks. Owing to 5-hmC levels being O_2-dependent (see section 4.3.2), and the fact that the JmjC-KDMs also contain a 2-OGDO domain, it was decided to examine the effects of low [O_2] on the global protein expression of these histone marks. Protein samples were harvested from undifferentiated and day 3-differentiated mESCs, exposed to atmospheric or 1% O_2, and were probed for H3K27me2/3 and H3K4me3 (Figure 4:29). Protein expression of the repressive marks (H3K27me2 and H3K27me3) were decreased from undifferentiated to day 3-differentiated mESCs in atmospheric O_2, whereas H3K4me3 was unaffected. Differentiation of mESCs in 1% O_2 prevented this loss of H3K27me2 and H3K27me3 at the onset of differentiation, while inducing H3K4me3 expression by approximately 4-fold above levels in undifferentiated (and day 3-differentiated) mESCs.

Bivalent promoters silence developmental genes in ES cells while keeping them poised for activation[459,461]. Therefore, the loss of the repressive marks (and maintenance of the activating mark) at the onset of cellular differentiation may be a characteristic trait. Furthermore, the Prc2 mediated H3K27me3 in mESCs is required for pluripotency[484,485]. Therefore, 1% O_2 appears to maintain the undifferentiated pluripotent state, consistent with the increase in pluripotent marker expression as shown in Figure 3:11 and Figure 3:12. The elevation of the activating mark H3K4me3 in 1% O_2 may account for increased Tet3 expression at day 3, (see Figure 4:28). However, more experiments would be needed to confirm the increase of this histone mark at the Tet3 promoter in low [O_2]. Despite this, it would be assumed this was a temporal effect, as Tet3 expression is reduced in 1% O_2 at later time points.

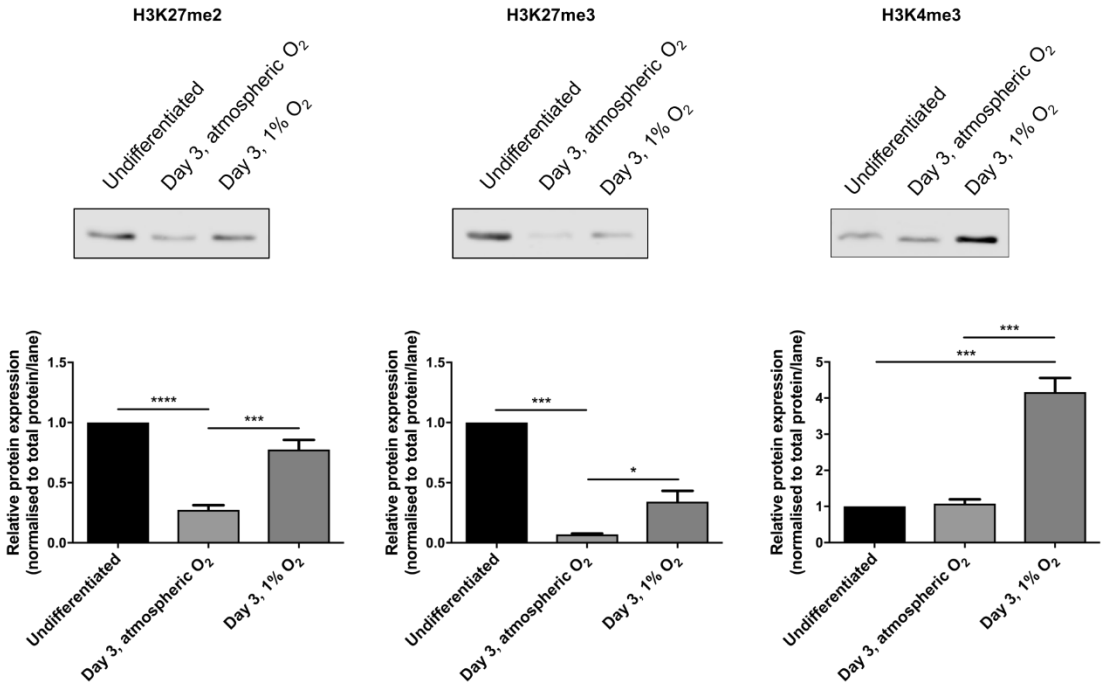


Figure 4:29 1% O_2 induced protein expression of H3K27me2/3 and H3K4me3 in day 3-differentiated mESCs. Protein was extracted from undifferentiated and day 3-differentiated mESCs exposed to atmospheric or 1% O_2. Protein expression of H3K27me2/3 and H3K4me3 was determined by Western blot analysis. Data are expressed as the mean ± SEM from 3 independent experiments, and expressed relative to shRNA control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. *p<0.05, *** p<0.001, **** p<0.0001

4.3.17 The mRNA expression of methyltransferase subunit of Prc2, Ezh2, is consistent with the loss and gain of H3K27me2/3 protein expression during differentiation under atmospheric and 1% O_2 respectively

The catalytic subunit of Prc2, Ezh2, is a histone methyltransferase that targets lysine-27 of histone H3[486]. To determine whether Ezh2 mRNA expression is consistent with the protein expression of H3K27me2/3 (Figure 4:29), Ezh2 levels were determined in undifferentiated and day 3-differentiated mESCs (exposed to atmospheric and 1% O_2) (Figure 4:30). Concomitant with the decrease in H3K27me2/3 protein, Ezh2 mRNA was decreased between undifferentiated and day 3-differentiated mESCs. Additionally, the increase in H3K27me2/3 protein after differentiation in 1% O_2 was also mirrored by increased Ezh2 expression. Thus, these data likely confirm Ezh2 is probably responsible for methylation of these histones in mESCs. Furthermore, the Ezh2 mRNA expression profile mimics pluripotent marker mRNA expression at the onset of differentiation (Figure 3:9) and upon exposure to 1% O_2 (Figure 3:11 and Figure 3:12), thereby potentially supporting the role of Ezh2/H3K27me3 in stem cell identity and pluripotency[487].

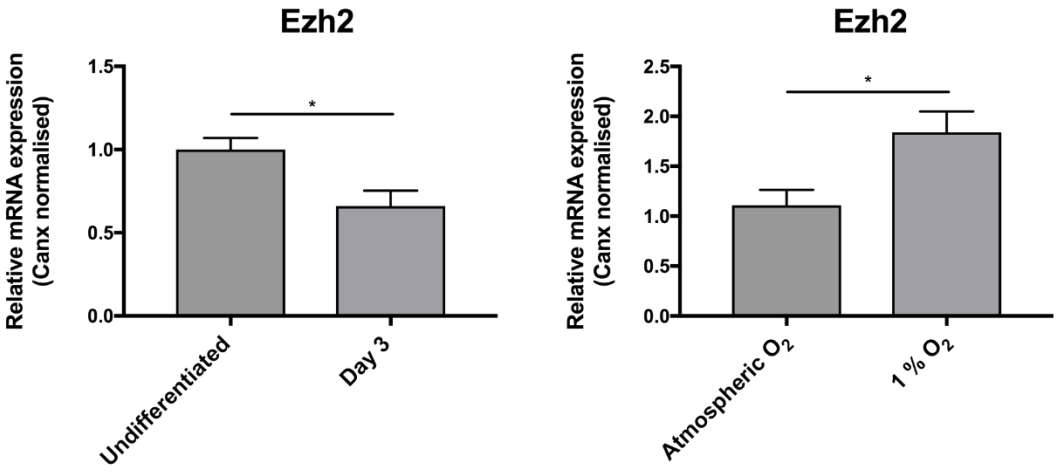


Figure 4:30 mRNA expression of Ezh2 is decreased between undifferentiated and day 3-differentiated mESCs, but is increased in differentiating cells when exposed to 1% O_2. Ezh2 mRNA expression was determined by QPCR in undifferentiated and day 3-differentiated mESCs exposed to atmospheric and 1% O_2. Data are expressed as the mean ± SEM relative to undifferentiated and day 3-differentiated atmospheric O_2 exposed mESCs. Data are analysed by an unpaired t-test. * $p<0.05$

4.3.18 Tet1 ablation in mESCs did not affect mRNA expression of Tet3 and its transcript variants during a mESC differentiation time course

It was established that low O_2 culture decreases Tet3 expression during mESC differentiation at day 7 and 11 time points (see Figure 4:28), which is consistent with a loss of 5-hmC from its promoter region. Furthermore, Tet1 was shown in part to be responsible for the burst of hydroxylation in day 3-differentiated mESCs, which was subsequently found to be predominantly targeted to this Tet3 promoter region. Based on this evidence it could be speculated that ablation of Tet1 would result in the loss of 5-hmC from the Tet3 promoter and would act to subsequently decrease Tet3 mRNA.

This hypothesis was tested by differentiating Tet1-ablated mESCs over a defined time course of differentiation, where the expression of Tet3 was previously found to be induced (between days 3 and 7) (Figure 4:5). The expression of Tet3 (denoted as Tet3 pan) and 2 known variants named Tet3 upstream (denoted Tet3 up) and Tet3 downstream (denoted Tet3 down) (see Chapter 5: Results 3 for further details) was determined (Figure 4:31). The expression of Tet3 transcript variants was investigated in parallel to confirm whether Tet1 may act to regulate the expression of a specific Tet3 paralogue. It should be noted here that the putative promoter region of the Tet3 downstream variant was the target of 5-hmC enrichment, as demonstrated from hMeDIP-seq analysis (Figure 4:19). Investigation into the regulation of expression of these 2 variants is the focus of Chapter 5: Results 3. Tet1-silencing failed to prevent the time-dependent increase in Tet3 pan mRNA expression (encompassing all variants) or of either the upstream or downstream variants. However, it was evident that compared to the relative expression in undifferentiated cells, Tet3 downstream mRNA showed less induction during differentiation than the upstream variant.

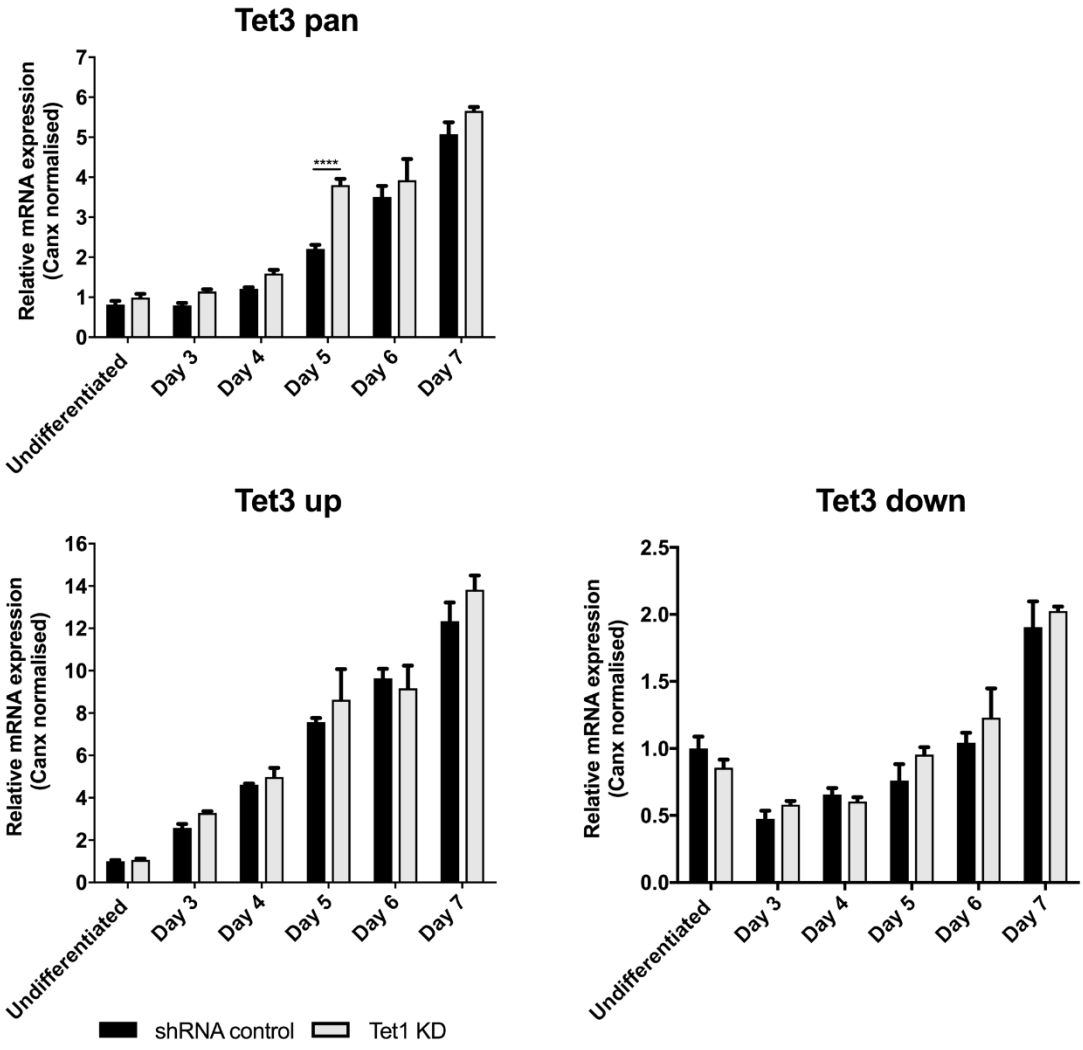


Figure 4:31 Tet1 albated mESCs had no affect on the mRNA expression of Tet3 pan and Tet3 up/downstream transcript variants. mRNA expression of pan Tet3 and both upstream and downstream variants over a 7 day time course of mESC differentiation was assessed by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated mESCs. Data are analysed by a two-way ANOVA with Tukey's post-hoc analysis. **** p<0.0001. KD: knockdown.

4.3.19 Tet1-ablated mESCs failed to reduce 5-hmC within the Tet3 promoter region

Tet1 silencing was unable to reduce Tet3 mRNA expression during mESC differentiation (Figure 4:31). To determine whether this was a result of having no effect on 5-hmC levels within the Tet3 promoter region, hMeDIP-PCR was conducted on Tet1-depleted mESCs, which had been differentiated to day 3 (Figure 4:32). It was apparent that Tet3 was enriched at day 3 in shRNA control cells, above IgG control and reduced in 1% O_2 (as also shown in Figure 4:27). However, Tet1-ablated mESCs failed to display reduced levels of 5-hmC. This result was also confirmed by bisulfite and oxidative bisulfite treatment (data not shown), as performed in Figure 4:21. It may be speculated that this surprising lack of effect of Tet1 silencing upon Tet3 expression (and hydroxylation of its promoter region) may be a result of functional redundancy, see Chapter 6: General Discussion.

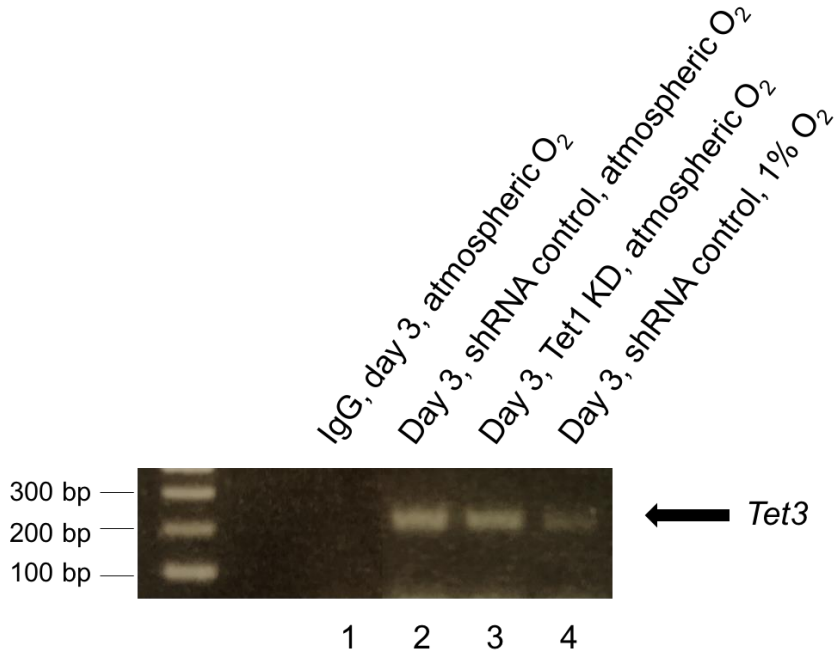


Figure 4:32 Tet1-ablated mESCs had no effect on 5-hmC levels within the Tet3 promoter region. shRNA control and Tet1-ablated mESCs were differentiated for 3 days (shRNA control cells were also differentiated for 3 days in 1% O_2) and genomic DNA isolated for hMeDIP-PCR. A Tet3 promoter fragment that encompassed the area of 5-hmC enrichment identifed from the hMeDIP-seq study was amplifed by PCR and visualised by agarose-gel electrophoresis. KD: knockdown.

4.4 *Discussion*

4.4.1 Assessment of DNA methylation marks during mESC differentiation revealed a novel transient burst of 5-hmC in day 3-differentiated cells

Examination of 5-hmC levels during an 11-day time course of mESC differentiation revealed a dynamic expression profile with a transient burst of hydroxylation at day 3 (Figure 4:1 and Figure 4:2). 5-hmC has previously been found to be highly expressed in ESCs[168,252], but as cells differentiate and become specialised (except within neuronal cells) levels decline[249]. Few studies have measured 5-hmC levels at the onset of differentiation, however some groups have recorded initial dynamic changes[422,423], which are perhaps consistent with the novel hydroxylation burst described here. The validity of this result was confirmed by 2 independent analyses; dot-blot and mass-spectrometry, which showed a high level of correlation. Additionally, the quantitation by mass-spectrometry of methylation marks in undifferentiated mESCs demonstrated 5-hmC is approximately 0.08% and 5-mC approximately 0.6% of total C, which is largely in agreement with a previously published study[168].

Surprisingly, the increase in 5-hmC from day 1 of mESC differentiation was not preceded by an elevation of 5-mC. Previously it has been reported 5-hmC depends on pre-existing 5-mC[186,488]. Therefore, this may suggest hydroxylation occurs at already existing methylated residues, or alternatively the methylation of C and the subsequent hydroxylation of 5-mC are tightly coupled at this stage of differentiation, so as to ensure a high rate of turnover resulting in immediate 5-hmC formation.

De novo generation of 5-mC occurs rapidly after day 1 of mESC differentiation (Figure 4:2), and plateaus by day 5. This may be consistent with methylation functioning to silence pluripotent gene transcription as the onset of differentiation[489]. Pluripotent genes, such as Oct4[490], Nanog and Lefty1[491] are protected from *de novo* methylation silencing in ESCs and are thus hypomethylated, however in response to differentiation signals acquire a repressed chromatin state that is thought to involve gain of 5-mC and loss of H3K4me3[492]. In addition, DNA methylation profiling demonstrated that early induction of differentiation from mESCs into 3 germs layers (as conducted in these experiments) resulted in gains of promoter methylation at germ cell-associated genes, which also correlated with the loss of H3K4me3, and is thus associated with a transcriptionally inactive state[493]. This therefore functions to restrict cell fate towards a somatic lineage[493]. However, DNA methylation has also been linked to the precise temporal and spatial control of tissue-specific genes. This was proposed from the genetic ablation of Dnmt

enzymes in mESCs, which revealed the upregulation 337 tissue-specific genes[494]. Hence DNA methylation is involved in gene repression and therefore DNA demethylation may function to activate tissue-specific gene transcription[494]. Thus, in combination with the strikingly different profiles of 5-mC and 5-hmC observed here during mESC differentiation, it is apparent that 5-hmC is not merely a consequence of differing 5-mC levels, but may itself have an independent functional role in differentiation events.

4.4.2 Tet1 is predominantly responsible for the O_2-dependent hydroxylation burst at day 3 of mESC differentiation

Differentiation of mESCs under low O_2 tensions revealed 1% O_2 had a more profound inhibitory effect on 5-hmC levels than 3% O_2 (Figure 4:3 and Figure 4:4). Given the requirement of 2-OGDO domain containing enzymes for O_2 it is perhaps unsurprising Tet activity is reduced, resulting in decreased hydroxylation. However, as shown with ectopic expression of Tet3 in HEK cells, 5-hmC levels were unexpectedly elevated in 3% O_2, which was alluded to be potentially caused by differential O_2-dependent regulation of cellular metabolites (Figure 3:25). This elevation of 5-hmC in 'hypoxia' is evident in the literature[282]. In neuroblastoma cells, 1% O_2 exposure induced 5-hmC by approximately 2/3-fold, which was found to be attributed to increased Tet1 expression[495]. Similarly, Tet1 and Tet3 expression was also elevated in 1% O_2 exposed primary isolated breast cancer cells and in some breast cancer cell lines[496]. A correlation between Tet expression and 5-hmC levels in a plethora of cell lines exposed to 0.5% O_2, revealed that cells without hypoxic induction of Tet mRNA resulted in reduced 5-hmC, whereas cells that showed upregulation of Tet mRNA could function to compensate and thus increase 5-hmC[309]. Intriguingly, within this study, undifferentiated mESCs were shown to have reduced 5-hmC levels upon exposure to 0.5% O_2[309], and therefore the Tet inhibitory potential in these cells is in agreement with data shown here. However, it can be noted no loss of 5-hmC was observed in the undifferentiated mESCs used here at the tested 3% and 1% O_2 (Figure 4:3 and Figure 4:4). Collectively, it is therefore apparent the relationship between O_2 and 5-hmC may be cell-type dependent.

Based on the assessment of Tet activity in HEK cells to differential physiological O_2 concentrations (Figure 3:23 and Figure 3:24), Tet1 was found to be the most likely paralogue inhibited by low O_2 tensions. Tet1 activity was significantly reduced at 3% O_2, however 5-hmC levels at this O_2 tension over the mESC differentiation time course were largely unaffected, compared to 1% O_2. This may be explained by the apparent increase in ectopically expressed Tet3 activity at 3% O_2 in HEK cells, which is also reflected by elevated 5-hmC levels in control-transfected cells at this O_2 tension (Figure 3:24). Hence,

the inhibitory effect of Tet1 activity at 3% O_2 may have been masked by increased Tet3 activity.

For Tet1 to be responsible for regulating the O_2-dependent burst in day 3-differentiated mESCs it must be expressed. Assessment of Tet1, Tet2, and Tet3 mRNA expression revealed that Tet1 (and Tet2) decreased while Tet3 increased over the 11-day time course of differentiation (Figure 4:5). This expression pattern, also reported previously[252], highlights a discrepancy within the data, such that the mRNA profile doesn't match elevated 5-hmC levels at day 3. Upon closer inspection of the relative expression levels of the Tets in day 3-differentiated mESCs, Tet1 was found to be the most abundant (Figure 4:6) and accordingly speculated to be predominant paralogue responsible for the hydroxylation burst. This was confirmed through differentiation of Tet-ablated mESCs (Figure 4:7 and Figure 4:8). Tet1 silencing did not result in total loss of 5-hmC, which may be accounted for by only partial knockdown of Tet1 and/or potential compensatory effects of Tet2 and Tet3 (see Chapter 6: General Discussion). In addition, considering Tet1 silencing and 1% O_2 reduced 5-hmC levels in day 3-differentiated mESCs to equivalent amounts (Figure 4:9), it could be inferred this burst of hydroxylation is mediated *via* Tet1 in an O_2-dependent manner. This Tet1-controlled burst of hydroxylation has not been reported previously in early differentiated ESCs. However, *in vivo* Tet1 has uniquely been shown to be responsible for hydroxylation in the Epi and thus contribute to the dynamic 5-hmC levels in the early developing mouse embryo[497].

4.4.3 Ascorbic Acid is not responsible for the increased catalytic activity of Tet enzymes at the onset of mESC differentiation

The discrepancy between Tet mRNA expression and the burst of hydroxylation in day 3-differentiated mESCs may suggest the catalytic capacity of Tet enzymes becomes increased. In addition to TCA cycle metabolites[303] and O_2 as regulators of Tet activity, AA has been shown to increase 5-hmC by reducing Fe^{3+} to Fe^{2+} and thereby maintaining the catalytic capacity of the enzymes[441]. During early ESC differentiation, AA levels have been shown to increase up to day 4[498], which coincides with the induction of 5-hmC at the onset of differentiation reported from this study (Figure 4:1 and Figure 4:2). Therefore, it could be speculated that AA may play a crucial role in increasing the catalytic activity of Tet1 to induce this burst of hydroxylation. Hence the ability of AA to promote 5-hmC generation *via* positive regulation of Tet activity was investigated. However, collectively the data attained here did not support a role for AA in hydroxylation regulation in differentiating cells. Importantly, it should be additionally noted here that the novel burst of hydroxylation was not an artefact of cell culture conditions (owing to the

different concentrations of AA contained within EmbryoMax® FBS and KO serum replacement) (Figure 4:17). This further supports the notion that the day 3 burst of hydroxylation is independent of AA. Therefore, the precise molecule mechanism(s) which underlie increased Tet activity, as Tet1/2 mRNA expression falls and before Tet3 expression increases, remains to be elucidated.

4.4.4 The transient O_2-dependent burst of hydroxylation in day 3-differentiated mESCs is predominantly targeted to the Tet3 promoter region

A surprising finding of this study was that the most significant target of this burst of hydroxylation in day 3-differentiated mESCs was a CG-rich Tet3 promoter region (Figure 4:19). It was further established that 5-hmC enrichment of this promoter occurred *de novo* after mESC were induced to differentiate, and thus is in accordance with the global 5-hmC induction seen at day 3 (Figure 4:21). Considering Tet1 was predominantly found to be responsible for this O_2-dependent burst of hydroxylation in day 3-differentiated mESCs (Figure 4:7 and Figure 4:8), it is plausible to believe that Tet1 is thus responsible for 5-hmC enrichment on the Tet3 promoter (and the B230319CO9Rik non-coding transcript). Consistent with this hypothesis, the reduction in global 5-hmC under 1% O_2 (Figure 4:3 and Figure 4:4), which could be attributed to the O_2-sensing capabilities of Tet1 (Figure 3:23 and Figure 3:24), also resulted in the loss of 5-hmC from the Tet3 promoter region (Figure 4:27).

Intriguingly, the hydroxylation on the Tet3 promoter at day 3 correlated with induction of Tet3 expression (between days 3 and 7) and similarly, the loss of hydroxylation in 1% O_2 resulted in decreased Tet3 expression (at days 7 and 11) (Figure 4:28). Perhaps interestingly Tet3 expression was also reduced at 3% O_2, but was found to not correlate with the loss of global 5-hmC during mESC differentiation under this O_2 tension. Thus, it could be deduced that this effect may be Tet1-dependent and site specific, as ectopically expressed Tet1 activity in HEK cells demonstrated inhibition at 3% O_2 (Figure 3:23 and Figure 3:24). Hence Tet1-mediated loss hydroxylation at this Tet3 promoter region in 3% O_2 may account for the reduced Tet3 mRNA expression. Currently there is no literature evidence suggesting a regulatory role for Tet1 in the transcriptional regulation of Tet3, except from some RNA-seq data from Tet1-silenced undifferentiated mESCs[323]. Thus it was hypothesised Tet1 silencing would reduce 5-hmC within the Tet3 promoter and subsequently blunt its mRNA expression.

Assessment of Tet3 expression, over the time points of induction (days 3-7) (Figure 4:5), revealed no difference between Tet1-silenced and shRNA control mESCs (Figure 4:31).

Upon closer inspection it was revealed that the 5-hmC enriched region within the Tet3 locus corresponded to a promoter region containing transcriptional start sites for an isoform that does not contain the CXXC DNA binding domain[175]. Expression analysis of this isoform, denoted Tet3 downstream, and the full length CXXC-containing variant, Tet3 upstream, was also determined between days 3 and 7, in both Tet1-ablated and shRNA control mESCs (Figure 4:31). Tet1-silencing had no effect upon either transcript variant, however interestingly the upstream transcript showed a greater mRNA induction relative to undifferentiated cells, than the downstream. Hence, it might be inferred that enrichment of 5-hmC could be functioning as a transcriptional repressor, which appears to contradict the belief 5-hmC is associated with euchromatin and increased transcription[186]. However, 5-hmC-mediated transcriptional regulation may be more complex, and in fact dependent on its genomic positioning, for example it has been suggested that promoter and gene body hydroxylation may contribute to gene repression and activation respectively[390].

Furthermore, from the hMeDIP-PCR study (Figure 4:32), Tet1-ablated mESCs showed no reduction in 5-hmC from the Tet3 promoter, which potentially may explain why no change in Tet3 mRNA expression was noted. This may be due to Tet1 being only partially silenced, meaning there may be sufficient levels to sustain hydroxylation (a small induction of 5-hmC above undifferentiated cells in day 3 is evident in Tet1 KD cells) (Figure 4:7 and Figure 4:8). Alternatively, there may have been functional compensation or functional redundancy effects with Tet2 at this time point (see Chapter 6: General discussion for more details on this phenomenon). Nonetheless, the data presented here supports the enrichment of hydroxylation on the downstream Tet3 promoter *via* O_2-dependent regulation of Tet1.

4.4.5 Tet3 is a bivalent promoter in association with Tet1

ChIP analysis confirmed Tet1 was bound to this Tet3 promoter region in both undifferentiated and day 3-differentiated mESCs (Figure 4:22). Tet1 exhibits a strong binding preference for unmethylated CG rich regions (or CpG islands) within the genome, which often overlaps with transcriptional start sites and therefore gene promoters[183,499]. Hence, due to the high CG content of the Tet3 downstream promoter, it is valid to reason that the binding of Tet1 *via* its N-terminal CXXC domain[500] can occur at this locus.

In the undifferentiated state Tet1 appears inactive, as 5-hmC is not detectable within this Tet3 promoter region (Figure 4:21). This may be a reflection of the low 5-mC levels

present in undifferentiated cells, as a result of the 2i culture conditions. 2i conditions have been shown to reduce DNA methylation, driven by a decrease in Dnmt enzyme expression, to a similar epigenetic pattern seen in the Epi of the pre-implantation embryo[501,502,503]. However, it should be noted that a level of methylation in these undifferentiated cells does exist, as treatment with AA induced 5-mC-hydroxylation (Figure 4:10). Furthermore, in agreement with the literature, the constitutively methylated CpG-poor promoter region of Orm1[216] was also found in the undifferentiated mESCs used in this study (*via* bisulfite and oxidative bisulfite treatment) to contain the 5-mC mark (but not the 5-hmC) (Data not shown). Importantly it should be noted that the induction of a naïve pluripotent state by culture of mESCs under 2i conditions does not alter their differentiation potential compared to serum (and LIF) maintained mESCs[462]. Overall, at the onset of differentiation, in the absence of 2i conditions, 5-mC levels rise, enabling Tet1 to induce hydroxylation at the Tet3 promoter. The presence of Tet1 at this Tet3 promoter potentially enables the close coupling of 5-mC to 5-hmC, reflected by 2 out of 3 tested CpGs at this promoter containing the 5-hmC mark (Figure 4:21).

In addition to Tet1, the Tet3 promoter region is also marked with the repressive (H3K27me3) and active (H3K4me3) histone marks, characteristic of a bivalent promoter. Bivalent promoters mark developmental genes, keeping them transcriptionally repressed in undifferentiated cells, but poised for immediate activation upon differentiation[459]. Consistent with this study, Tet3 is lowly expressed in undifferentiated cells, but becomes induced after day 3 of differentiation, correlating with the loss of the H3K27me3 mark and maintenance/gain of H3K4me3 (Figure 4:24 and Figure 4:25). To support the credibility of this finding, it has been reported that Tet1 and 5-hmC are enriched at gene promoters associated with bivalent domains[183,389,184]. Further, collated H3K4me3 and H3K27me3 ChIP-seq data sets from h/mESCs identified Tet1 as being highly enriched at bivalent domains[504]. This therefore potentially demonstrates Tet1 as a fundamental regulator of the balance between the maintenance of pluripotency and the onset of differentiation, and consequently therefore suggests that Tet3 is involved in developmental signalling. Interestingly, these domains are not enriched with 5-mC[494], which is consistent with the appearance of predominantly 5-hmC within this region in day 3-differentiated cells (Figure 4:21). This also highlights a distinct role between 5-mC and 5-hmC marks in transcriptional regulation.

4.4.6 Tet1 and 1% O_2 regulates the protein expression of repressive and active histone methylation marks

To gain a further mechanistic insight between Tets and histone modifications, the global protein expression of repressive and active histone marks was assessed in day 3-differentiated mESCs ablated of Tet1, Tet2 and Tet3 (Figure 4:26). The protein expression of H3K27me2 was also examined due to a recent study suggesting that the ratio between di- and tri- methylated forms of H3K27 can influence the differentiation ability of ESCs[505]. Ablation of Tet1, but not Tet2 or Tet3, induced protein expression of H3K27me2 and H3K4me3, which signifies the potential non-redundant functions of Tet proteins in early differentiating mESCs. The functional significance of this result is currently unclear, but further supports crosstalk specifically between Tet1 with histone methylation marks. Perhaps given the literature evidence it is not surprising that Tet1 interacts with histone marks[456,390,506], but considering depletion of Tet1 has previously been shown to disrupt Ezh2/Prc2 recruitment[194], it was perhaps surprising to see an induction of H3K27me2. However, this discrepancy may have occurred since protein expression data presented here was obtained from differentiated cells, whereas findings in the published study were from undifferentiated cells.

Assessment of the global protein expression of these histone methylation marks at the onset of differentiation demonstrated decreased expression of H3K27me2/3, while the levels H3K4me3 remained unchanged, compared to undifferentiated mESCs (Figure 4:29). This is consistent with the loss of a repressive signal, indicative of developmental gene activation. To determine whether [O_2] can also be a regulator of the histone methylation marks, protein expression was assessed after 3 days of differentiation in 1% O_2 (Figure 4:29). Low O_2 culture significantly elevated levels of both the repressive and activating marks, which potentially owing to their regulation of lineage control genes[507] may contribute to the skewed differentiation reported in Figure 3:11. Interestingly, Tet1-ablation and 1% O_2 induced similar inductive effects upon histone marks (except for H3K27me3), which may potentially further support a role of Tet1 in [O_2]-dependent signalling. The differential effect of Tet1-silencing on H3K27me3 and H3K4me3 is in agreement with a previous study highlighting the weak and strong correlations of Tet1 with these marks respectively[183].

Furthermore, the activity of the JmjC-KDMs are also potentially sensitive to changes in O_2 tensions owing to their 2-OGDO domain[267]. Previous studies have highlighted that histone methylation marks, including H3K4me3 and H3K27me3 are induced in low [O_2][508,509,510]. Specifically, H3K4me3 induction has been described to be more sensitive

to low O_2 tensions, compared to H3K27me3[511]. However, from the data displayed here (Figure 4:29), both marks show a similar fold induction at day 3 in 1% O_2, compared to atmospheric O_2. The increase in methylation marks has been described to be attributed to inhibition of JmjC-KDMs in low [O_2] conditions, however given most KDMs can target several histone residues (KDM5A-D targets H3K4me2/3 and KDM6A/B targets H3K27me2/3), the sensitivity of specific enzymes to O_2 require elucidation[511]. Despite this, it has been shown directly that inhibition of KDM5A in 1% O_2 can lead to an increase in global H3K4me3[509]. Additionally, it should also be considered that O_2 can regulate JmjC-KDMs (including the KDM5 and KDM6 families) transcriptionally[511]. Further to this, it has been identified that KDM5C and KDM6B have been demonstrated to be direct targets of HIF-1α[512,270]. Thus, similarly to as reported for the Tet enzymes, the action of O_2 to inhibit activity or increase transcription of JmjC-KDMs may be cell-type dependent[309]. Collectively, these data demonstrate cross-talk between O_2 and 2 independent epigenetic mechanisms, however how these may function to orchestrate transcriptional regulation requires further experimentation.

4.4.7 Conclusion

To summarise, the data shown here demonstrates that 5-hmC is dynamically expressed during mESC differentiation. This expression profile exhibits an early burst of O_2-dependent hydroxylation mediated *via* the catalytic activity of Tet1, which targets a promoter region spanning the transcriptional start site of a Tet3 isoform lacking a CXXC binding domain. This promoter region displays histone methylation marks, characteristic of a bivalent promoter and thus highlights Tet3 as a developmental and/or lineage-specific gene.

Chapter 5: Results 3; *Characterisation of the Tet3 Promoter(s)*

5.1 *Introduction*

5.1.1 Cellular specificity of Tet3

Tet3 mRNA expression was found to be induced during mESC differentiation (Figure 4:5), in agreement with the literature[386], suggesting Tet3 is likely to be implicated in cellular differentiation. This is supported from *in vivo* studies whereby the dual loss of Tet1 and Tet2 support the development of viable mice (albeit with a large fraction exhibiting perinatal lethality)[248], yet the combined loss of all 3 Tet enzymes contributed poorly to chimeric embryos and restricted developmental potential of EBs[254]. Independently, homozygous mutations of Tet3 has also been shown to result in lethality[392]. Specifically, the loss of Tet3 from ESCs, as mentioned in section 3.1.3, impairs neuronal differentiation *in vitro*[321]. A neurogenesis role of Tet3 has also been demonstrated through overexpression studies in fibroblasts cells, whereby Tet3 was shown to contribute to their reprogramming into functioning neurons[513]. This involvement of Tet3 in neurogenesis may be supported from expression data in the brain, whereby elevated Tet3 mRNA, compared to Tet2 and Tet1 was recorded in the cerebellum, cortex and hippocampus[422]. In addition, levels of 5-hmC have been recorded to be highest in the brain and CNS, compared to, for example the kidney, heart, lung and liver (whereas 5-mC showed little variation between such tissues)[260,514]. However, it is worth noting Tet1 and Tet2 have to some extent also been implicated in neuronal differentiation[261], potentially suggesting these enzymes function co-operatively to orchestrate neurogenesis (see Chapter 6: General Discussion).

Despite the involvement of Tet3 in neurogenesis and its prevalent expression in the brain, little is known about how Tet3 is transcriptionally regulated. Evidence from Chapter 4: Results 2 has suggested that Tet3 may be under epigenetic control *via* Tet1 activity, however mechanisms governing this apparent cellular specificity remains to be elucidated. A recent study functionally demonstrated Tet3 controls ectodermal and mesodermal cell fate decisions[322]. Ablation of Tet3 skewed mESC differentiation away from the neuroectoderm and towards the cardiac mesoderm lineage, whereas the reverse effect was true for ectopic Tet3 expression[322]. This suggests the precise transcriptional control of the Tet3 gene may be an important factor in the regulation of both neuronal and cardiac differentiation, however mechanisms that underlie this cellular specification remain unknown. Elucidation of such Tet3 regulatory mechanisms could enable improvements in neural-directed cellular differentiation, which may aid the development of cellular therapies to treat incurable neurodegenerative diseases, such as Alzheimer's, Huntingdon's, and Parkinson's. Conversely, inhibition of Tet3 may

function to promote cardiac differentiation, which may also aid development of cellular therapies to repair damaged myocardium after insult.

5.1.2 Tet3 has two distinct transcript variants

To investigate the regulation of Tet3 it is important to consider alternative protein isoforms, which themselves may have independent functions. At present, the Tet3 gene has been identified from a murine neural stem cell line to possess 2 major protein-coding transcript variants that differ in their N-terminal sequence[515]. The longer isoform contains a DNA-binding CXXC domain, while the other shorter variant does not (a distinct third isoform has been identified exclusively in mouse oocytes) (schematically shown in Figure 5:1)[175]. These transcripts originate from distinct promoter regions, suggesting that their expression may be regulated independently. Literature evidence has shown that both variants were lowly expressed in ESCs, but mRNA was induced once directed to undergo neural differentiation[515,175]. It can be noted that the CXXC-containing variant demonstrated a higher fold induction, relative to its shorter form[175], although the opposite effect was true from an earlier study[515]. The longer protein isoform may demonstrate neural specificity through selective binding to unmethylated CG promoters of neuronal-related factors, such as Pax6, Ngn2 and Tubb2b *via* its CXXC domain[176]. In comparison, the shorter variant, which was found to the be most abundant Tet3 isoform in the mouse retina, was targeted to DNA through an alternative mechanism[516]. In this case, the TF Rest targeted Tet3 to methylated silenced neuronal genes during retinal maturation to subsequently de-methylate and induce transcriptional activation[516]. The exact neuronal functions of both Tet3 isoforms remain controversial[515,175,516], however the current data do suggest that distinct Tet3 isoforms may have differing functional roles. It can be noted that from the data obtained in the unbiased mESC differentiation studies shown in Chapter 4: Results 2, 5-hmC was enriched specifically at the Tet3 promoter region that drives expression of the smaller, truncated, protein isoform (Figure 4:19).

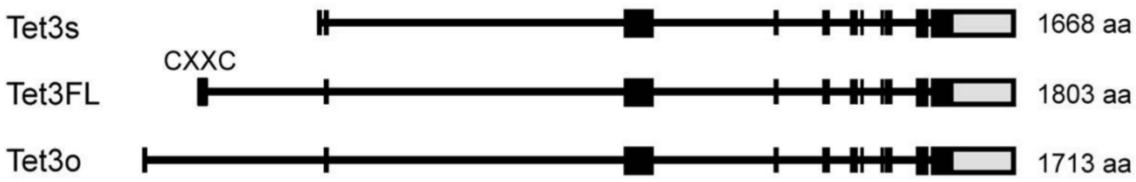


Figure 5:1 Schematic representation of Tet3 protein isoforms. Tet3 has been identified to have 3 distinct protein isoforms, denoted in this figure as Tet3-short (Tet3s), Tet3-full-length (Tet3FL) (containing an N-terminal CXXC domain) and a mouse oocyte specific variant named Tet3o. These are encoded by transcripts which comprise distinct non-coding upstream exons, suggesting they may be differentially regulated. Schematic taken from Jin, S *et al*. Cell Rep. 2016; 14(3):493-505[175].

5.2 *Aims*

To further elucidate the tissue specificity of the Tet enzymes and 5-hmC with respect to neural and cardiac tissues. Studies then aim to investigate the regulation of Tet3 that underlies cell specific function(s). To evaluate this the Tet3 transcript variants in unbiased differentiated mESCs will be defined and subsequently their distinct promoter regions will be characterised.

5.3 *Results*

5.3.1 Brain has higher levels of 5-hmC than heart in both mouse embryos and 1-week old pups

5-hmC has distinct tissue-specific patterns of expression, and is particularly found with high abundance in the brain. To verify the tissue specificity of 5-hmC, genomic DNA was isolated from brain and heart in murine E 18.5 embryos and 1-week old pups. 5-hmC and 5-mC levels were determined by dot-blot analysis (Figure 5:2 and Figure 5:3). Both embryos and 1-week old pups showed an equivalent ~2.5-fold elevation in the 5-hmC:5-mC ratio in brain, compared to heart. Mass-spectrometry was used to validate dot-blot data in 1-week old pup samples (Figure 5:4). These data clearly demonstrated increased 5-hmC in brain, compared to heart tissue, while 5-mC levels remained unchanged.

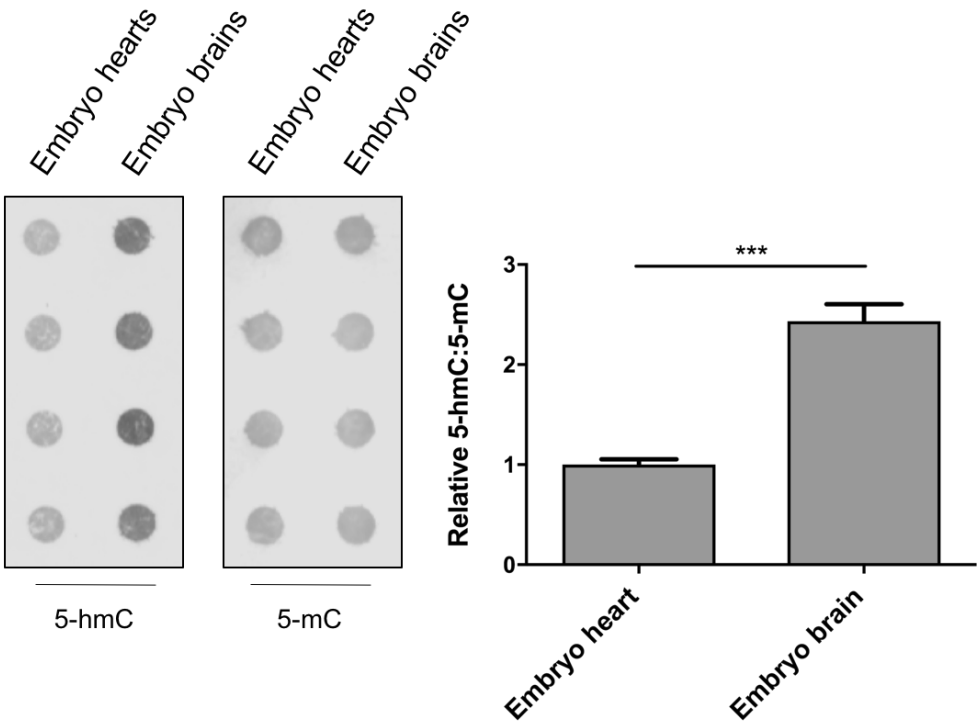


Figure 5:2 5-hmC levels are higher in mouse embryos brain than heart. Genomic DNA was isolated from heart and brain in mouse E 18.5 embryos and analysed for 5-hmC and 5-mC levels by dot-blot analysis. Blots were quantified and data displayed as the ratio of 5-hmC:5-mC. Data is expressed as the mean ± SEM and analysed by an unpaired t-test. *** $p<0.001$. n=4.

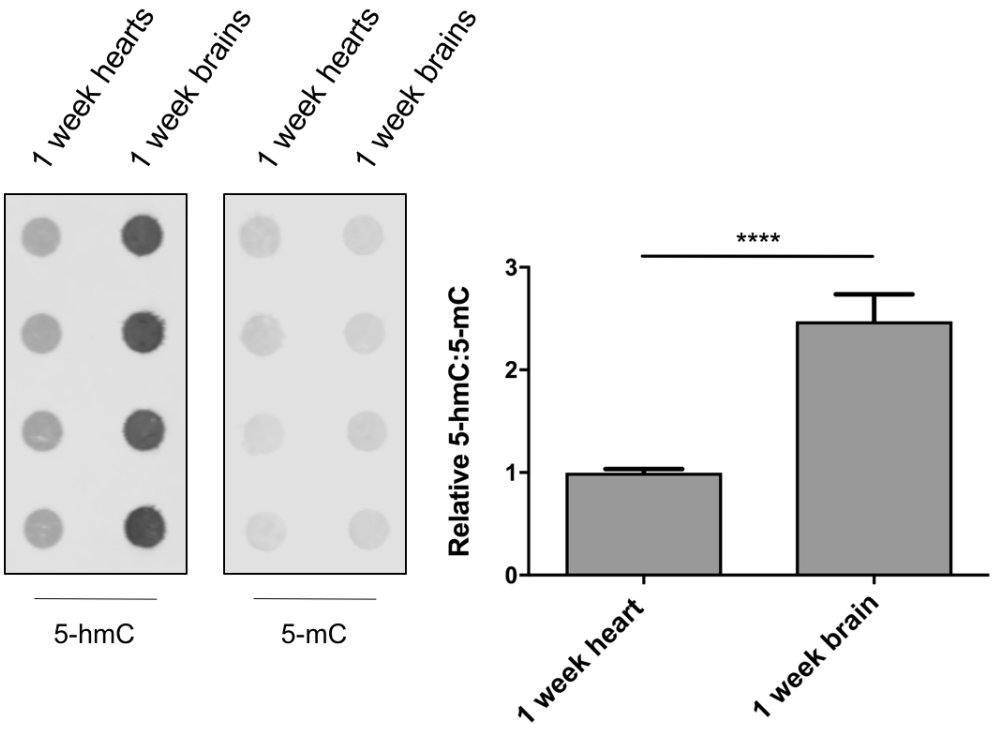


Figure 5:3 5-hmC levels are higher in 1-week old mouse pup brain than heart. Genomic DNA was isolated from heart and brain in 1-week old mouse pups and analysed for 5-hmC and 5-mC levels by dot-blot analysis. Blots were quantified and data displayed as the ratio of 5-hmC:5-mC. Data is expressed as the mean ± SEM and analysed by an unpaired t-test. **** $p<0.0001$. n=4.

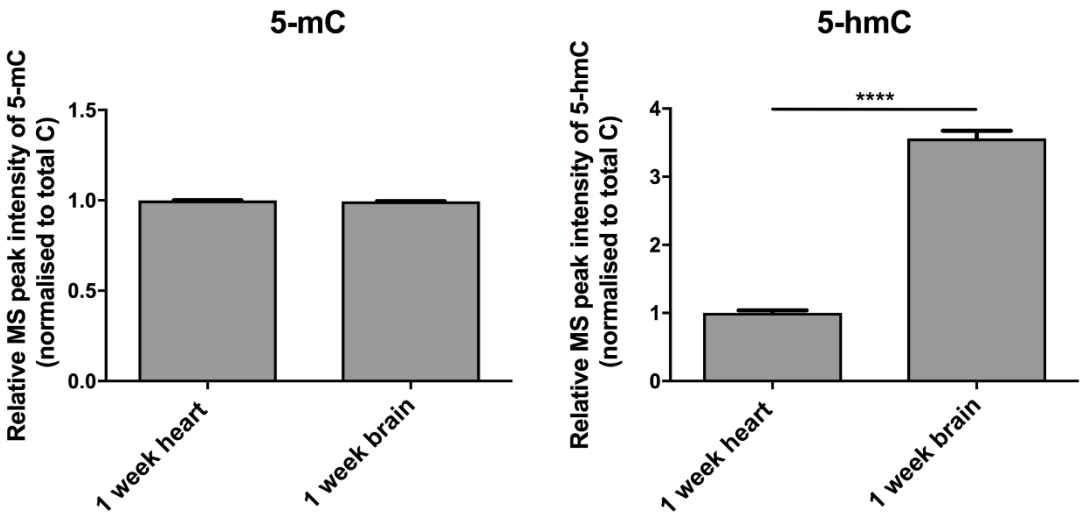


Figure 5:4 Mass-spectrometry analysis confirmed greater 5-hmC levels in 1-week old mouse pup brain compared to heart. Genomic DNA was isolated from heart and brain in 1-week old mouse pups and analysed for C, 5-mC and 5-hmC levels by mass-spectrometry analysis. Data are normalised to total C and displayed as the mean ± SEM. Data are analysed by an unpaired t-test. **** $p<0.0001$. n=4.

5.3.2 Tet3 mRNA expression accounts for elevated 5-hmC detected in mouse brain compared to heart

To investigate which Tet enzyme may be responsible for elevated 5-hmC in brain, compared to heart, the mRNA expression of each Tet paralogue was assessed in embryonic (E 18.5) (Figure 5:5) and 1-week old pup tissue (Figure 5:6). In the embryo no differences in Tet expression between heart and brain were detected. However, Tet3 trended towards an increased expression in brain, narrowly falling short of significance ($p=0.055$). In 1-week old pups Tet3 mRNA expression was significantly elevated in brain compared to heart. Tet1 mRNA was found to be decreased and Tet2 unaffected. Overall these data suggest Tet3 as the predominant Tet responsible for elevated 5-hmC levels in brain tissue. Note, despite protein detection as a more biological relevant readout accurate determination of protein expression was unattainable, attributed to the large size of the Tet proteins and difficulties accurately determining the correct protein bands.

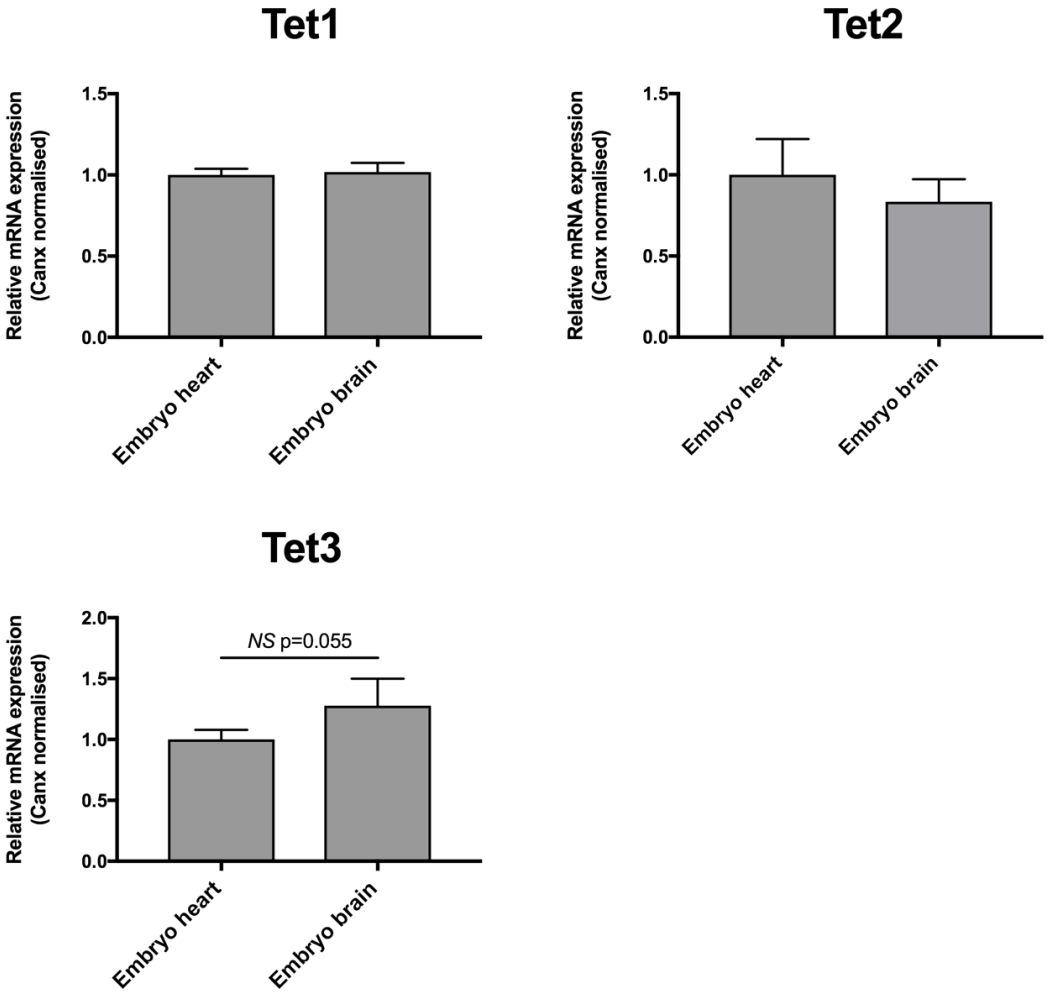


Figure 5:5 Tet3 mRNA expression trended to be elevated in mouse embryonic brain compared to heart. Tet1, 2 and 3 mRNA expression was determined by QPCR analysis from heart and brain in mouse E 18.5 embryos. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. n=4.

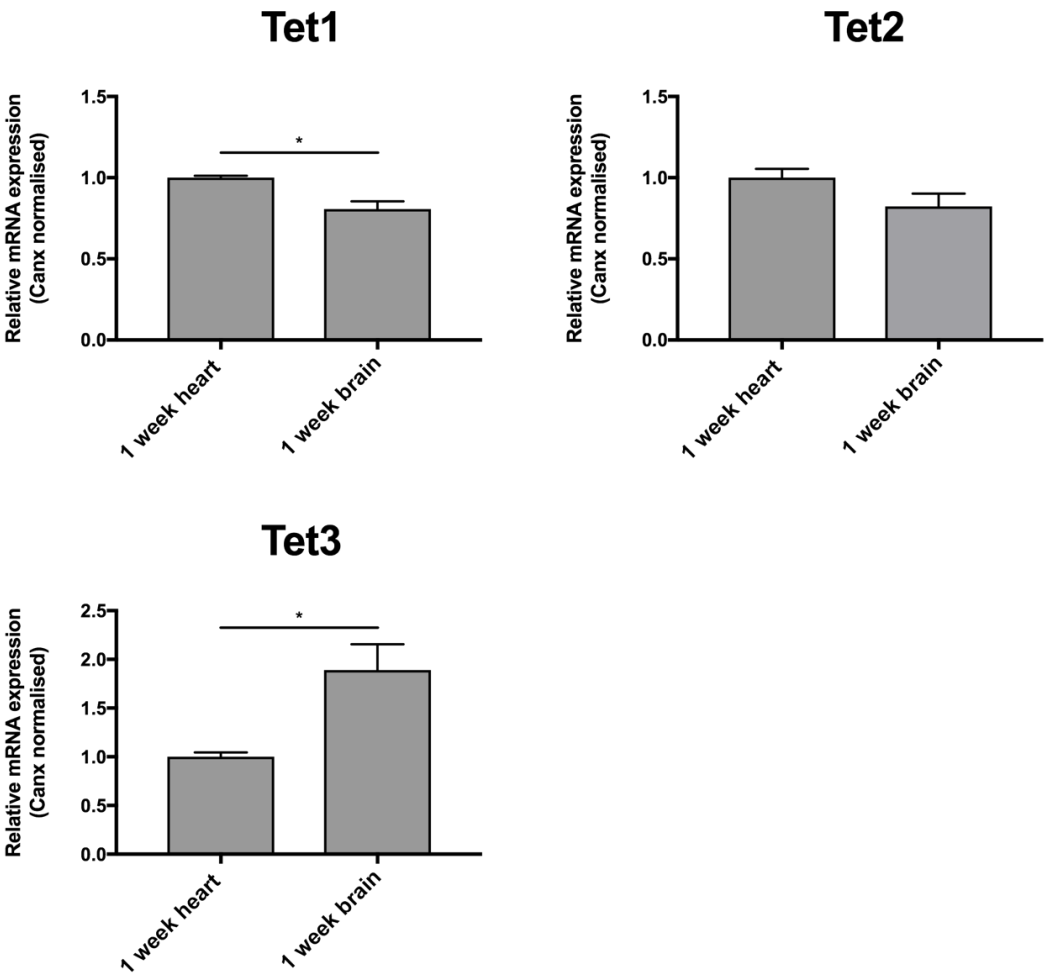


Figure 5:6 Tet3 mRNA expression is increased in 1-week old mouse pup brain compared to heart. Tet1, 2 and 3 mRNA expression was determined by QPCR analysis from heart and brain in mouse 1-week old pups. Data are expressed as the mean ± SEM and analysed by an unpaired t-test. *p<0.05. n=3.

5.3.3 Tet expression is induced in neural, compared to cardiac-directed P19 cell differentiation

To investigate the cellular specificity of Tet enzymes, P19 cells were directed to differentiate towards a neural or cardiac-like lineage in the presence of RA or DMSO respectively. To confirm successful induction of a neural and cardiac-like enriched cell fate mRNA expression of lineage-specific genes over a 10-day time course of differentiation was determined.

Greater induction of ectodermal markers was evident in RA-, compared to DMSO-treated differentiating P19 cells (Figure 5:7). β-III tubulin[343] trended towards increased expression at all tested time points in RA above DMSO treatment, and reached significance at days 4, 6 and 8. Dcx[354] expression was significantly induced at all time points in the presence of RA, and demonstrated a large induction that peaked at day 6 to approximately 450-fold above levels in undifferentiated P19 cells. In comparison, DMSO treated cells showed equivalent Dcx expression to undifferentiated cells at this time point. Ncam1[344] expression was induced with RA treatment after day 4 of differentiation, compared to undifferentiated P19 cells. Ncam1 mediates cell adhesions to modulate, for example, neuronal migration, neurite extension and synapse formation[517] and thus its expression profile shown here may reflect an involvement in the latter stages of neurogenesis. By contrast, mesodermal markers (Figure 5:8), in the presence of DMSO, were induced above RA treated cells at early differentiation time points. Gata4[342], Hand1[339] and Brachyury[338] were significantly induced above RA treated cells until after days 2, 4 and 6 respectively.

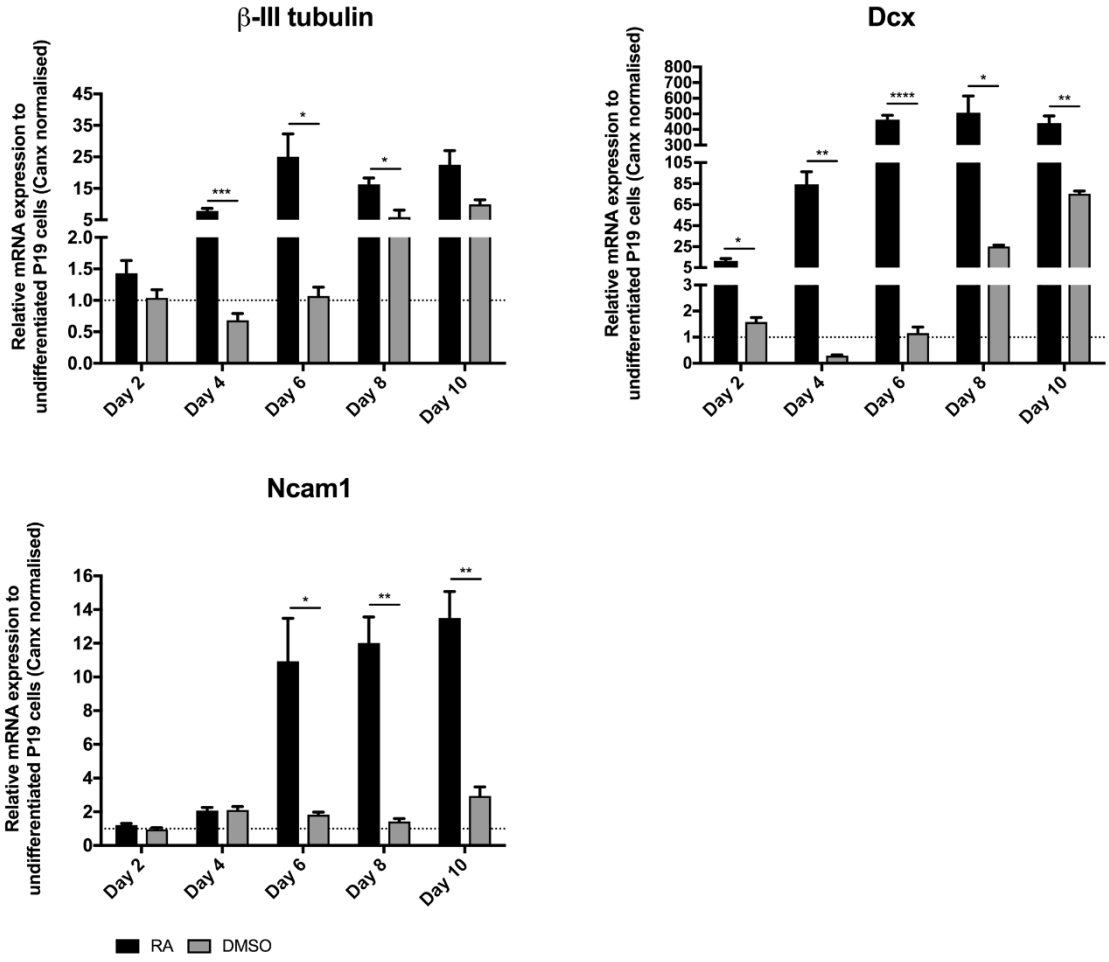


Figure 5:7 Ectodermal marker mRNA expression is induced with retinoic acid above dimethylsulfoxide treated differentiating P19 cells. P19 cells were induced to differentiate into embryoid body-like aggregates for 10-days in the presence of 1 μM RA and 1% DMSO. Ectodermal marker expression was assessed by QPCR analysis. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by multiple t-tests between RA and DMSO treated cells at each time point. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

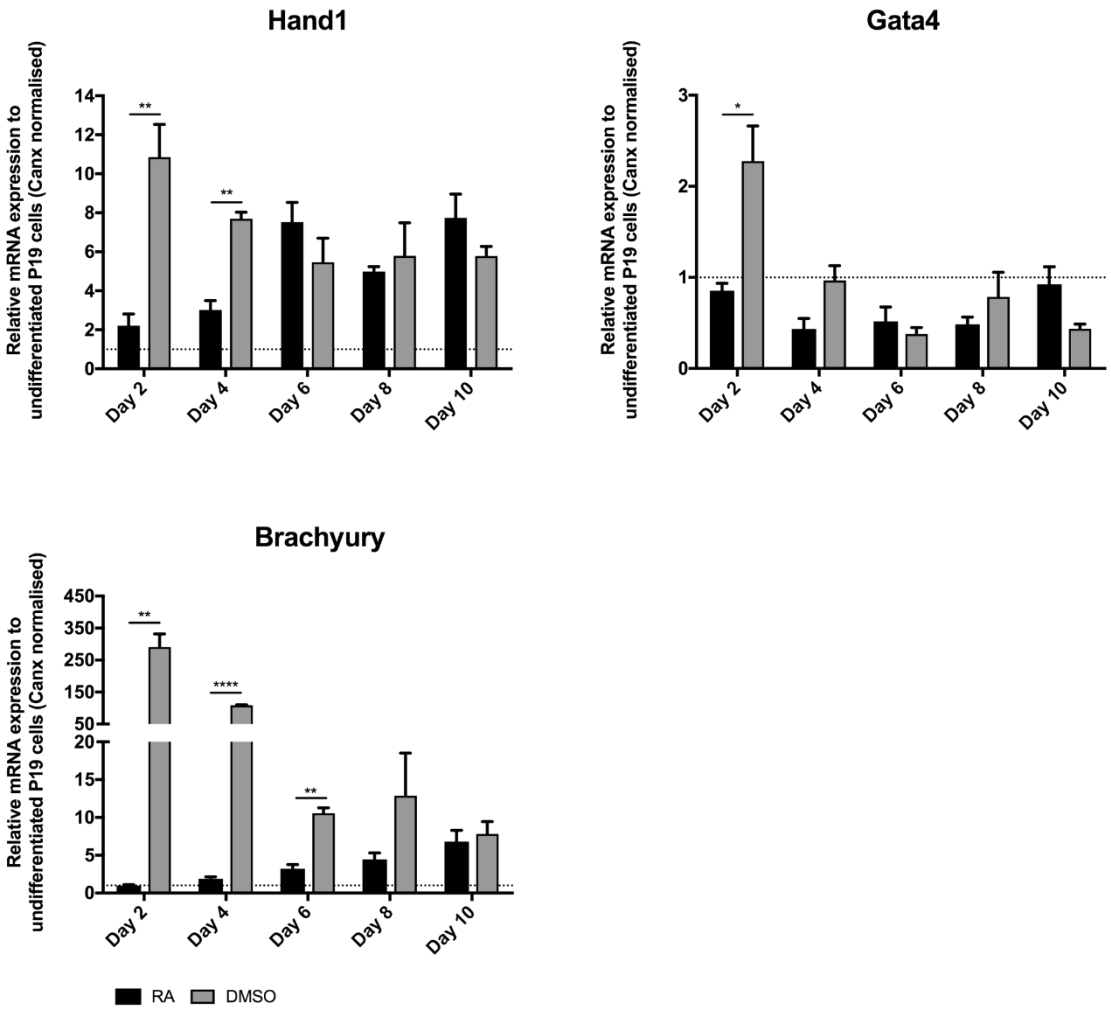


Figure 5:8 Mesodermal marker mRNA expression is induced with dimethyl sulfoxide above retinoic acid treated differentiating P19 cells. P19 cells were induced to differentiate into embryoid body-like aggregates for 10-days in the presence of 1 μM RA and 1% DMSO. Mesodermal marker expression was assessed by QPCR analysis. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by multiple t-tests between RA and DMSO treated cells at each time point. * $p<0.05$, ** $p<0.01$, **** $p<0.0001$

Thus RA and DMSO treatment could induce P19 cells to differentiate into cell populations characteristic of a neural- and cardiac-like lineage respectively. To investigate whether Tet enzymes are differentially associated with these 2 cell populations, perhaps consistent with them having an intrinsic role in cell specification, mRNA expression of Tet1, Tet2 and Tet3 was determined over the RA- and DMSO-treated P19 cell differentiation time course (Figure 5:9). Overall, Tet enzymes appear more highly expressed in RA, compared to DMSO treated P19 cells. Tet1 mRNA appeared to be reduced or unchanged relative to undifferentiated cells (depicted by the line at 1) after either RA or DMSO treatment, perhaps consistent with the involvement of Tet1 in pluripotency regulation[252,506]. The expression levels of Tet2 and Tet3 in RA-treated cells were induced above their relative levels in undifferentiated P19 cells after 2 days of differentiation, and consistently elevated above their expression levels in DMSO-treated cells over this time course. Tet2 and Tet3 mRNA peaked at day 6 with RA treatment. However, DMSO induced a marked increase in Tet2 expression compared to undifferentiated P19 cells, but did not affect Tet3 mRNA expression.

These data suggest Tets may have a greater role in ectodermal, compared to mesodermal differentiation. This is in agreement with the elevated 5-hmC levels detected in brain, compared to heart tissue (Figure 5:3). However, contrary to the mRNA expression in tissue, in addition to induced Tet3 expression, Tet2 (and also to a lesser extent Tet1) was also found to be induced in RA- compared to DMSO-treated P19 cells. This may thus highlight interplay between the Tet enzymes. However, based upon the breadth of literature linking Tet3 function to cellular specificity[321,506,518], the focus of future studies in this chapter are on Tet3.

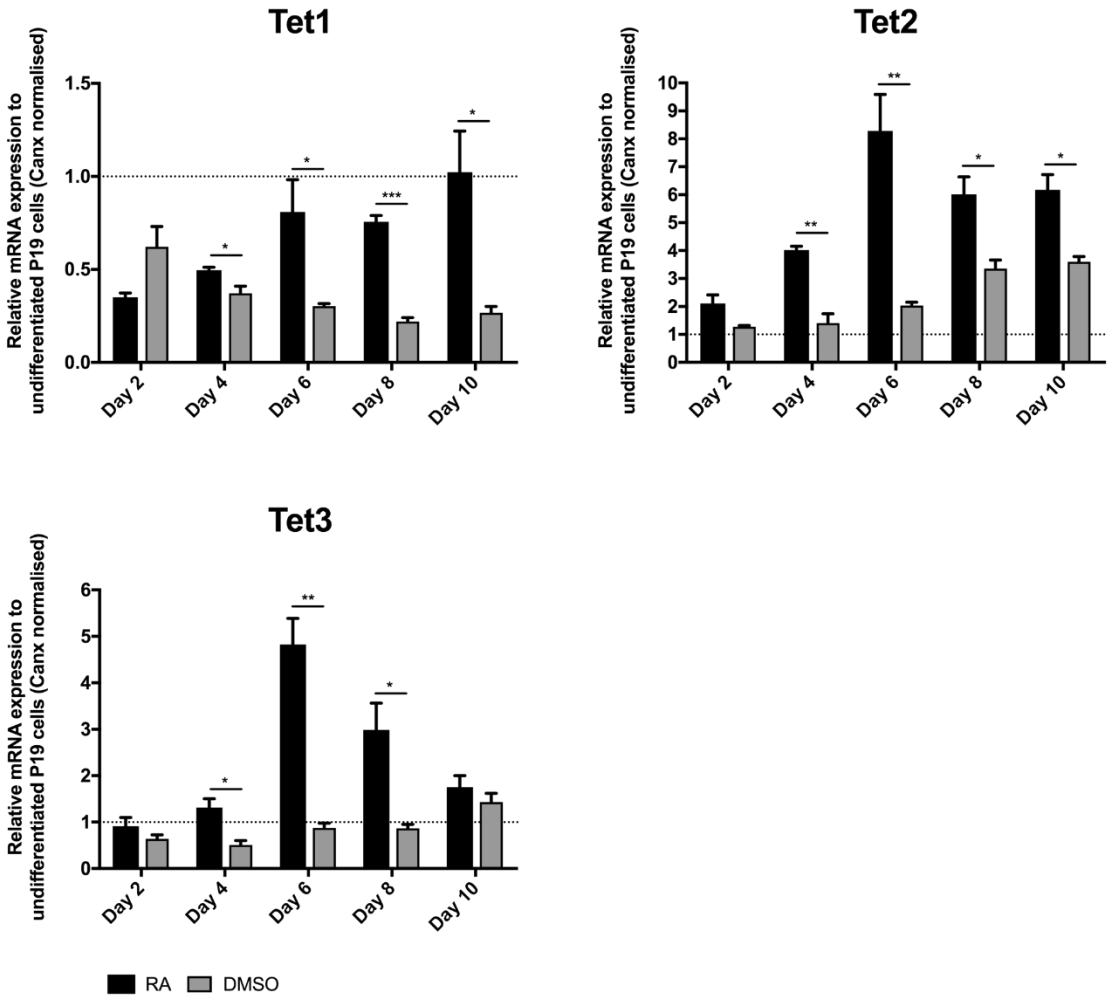


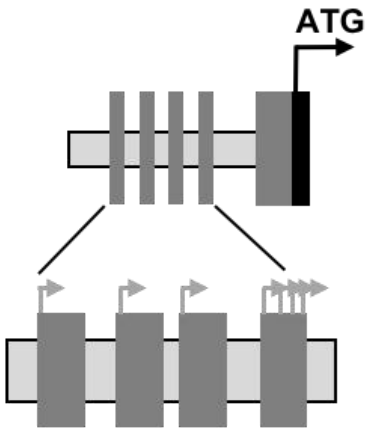
Figure 5:9 Tet enzyme expression is induced with retinoic acid above dimethyl sulfoxide treated differentiating P19 cells. P19 cells were induced to differentiate into embryoid body-like aggregates for 10 days in the presence of 1 μM RA and 1% DMSO. Tet1, 2, and 3 expression was assessed by QPCR analysis. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by multiple t-tests between RA and DMSO treated cells at each time point. * $p<0.05$, ** $p<0.01$, *** $p<0.001$

5.3.4 Tet3 has multiple transcriptional starts that encodes two distinct protein isoforms in day 7-differentiated mESCs

Elevated 5-hmC levels in brain, compared to heart tissue (see section 5.3.1), was predominately attributed to Tet3 mRNA expression (see section 5.3.2). To gain an insight into the transcriptional regulation of Tet3 underpinning tissue-specific effects, 5'RACE was performed to identify transcriptional start sites and hence promoter regions.

5'RACE was performed on mRNA taken from day 7-differentiated mESCs, the time point at which Tet3 expression was found to be elevated (Figure 4:5). Sequencing the clones of 5'RACE products revealed multiple cap sites that were found to be positioned upstream of 2 start codons (ATG), which encode 2 distinct protein isoforms. In combination with Tet3 sequence data available from Ensemble (https://www.ensembl.org) and NCBI (https://www.ncbi.nlm.nih.gov), multiple cap sites within 3 distinct, previously observed, non-coding exons (NCBI accession numbers: NM_001347313.1, XM_006505777.3, XM_0065057776.3, XM_006505775.3 and XM_017321465.1) and an additional novel non-coding exon (the most 5' non-coding exon) were confirmed to be associated with the upstream start site (Figure 5:10A). In addition, multiple cap sites were confirmed in the single non-coding exon at the more downstream start site (NCBI accession: NM_183138.2) (Figure 5:10B). These transcripts that give rise to 2 distinct protein isoforms are named here as Tet3 upstream (NCBI accession numbers: NP_001334242.1, XP_017176954.1, XP_006505840.1, XP_006505839.1, XP_006505838.1) and Tet3 downstream (NCBI accession number: NP_898961.2), marked in Figure 5:11. The Tet3 upstream transcripts encode a longer protein sequence that contains a CXXC DNA binding domain, which is absent in the shorter Tet3 downstream encoded protein (previously termed Tet3FL and Tet3s respectively by Jin, S *et al* (2016)[175]). These distinct proteins may therefore have separate functions, and thus the study of their regulation may allude to differential roles in tissue specification.

A



CCAGGTACCTGGCTGGGTGAGCACCGCCCAAACTGGTCGGGCTAGTTCACCCAGTGTCTG
TTTCTGCCATGACTTGCCCTTGTGTAGGGAGGGTACAGACCTACCCTTCTCCACAGCAAC
CCGGTAGGTGGAAGGCAGCTGTGGGCATTAGTCTGGCAGTCAGATCAGAAGGGTGGAGCT
GAATTAATCTAACCTGGAATTTCAACTTGGCCAGTAACAACAGGTGAGGGATGGGGGAGA
ATTAGCACAGAGAAATGGACAGTTTGGCAAGTTGGAAATACTCTTATGTCTTGTCCCTTA
CCGCCTGCCTAGAACACAGCCCAGGGTGACTTTTTCAAAGTGTCTGCATTATTTATGCCT
CGGTGAGGACTTGGGCCCTCGGTGACTGAGCCCATCCCGAATCTGCCTGGGTTCCTCTCT
GCGACGCGCCCCTTCCCTCTCCTCTGGGAGGCAGGGCGGGGGTCTCGGACTAGTCTTTGT
TTTAGATGTGCAACAGAGCCCGAGAGCAGCGCCTGGAAGGCTGGCGATGGCAGAGCCGTA
GGGTCGGCCAGGGCAGGGCGCTGGGTTGTGACCCTCATGCGGATGCCGTGCAGCGGGCGC
TCGGTCTCTGCCCTCGGGCACCCACTTCTTGCCGGTCAGGCTGACTCGGTACGCAAACGC
GCGATGTCCCCTATTGTCCTGCCTCGGCCTGAGAGGCAGCCGAGGAGGAGAAAGGGGGTC
GGTCGGGCTCCGGGGCAGTCGGGCGAGCGTGGGGCCGTCCCGAGCTCGCAGGGAGTGGCA
GGAAGTGGGGACACCCAGGCGGCTGGGCCAAGGGCCTAGCTGGGGAGGTCCGCGCCCCGC
CCCGCCCCTCCTCGACAGCTTGGGCTGCGGGCCCCGCCGGAGGGCAGGGAAAGGGCCAGC
CCGCGGGGACAGCCGCGTCTCAGGAGGGCCCCTCCGTGAGGCCTGGGGAGGCGCTGCAGG
TGCAAAACCCGGGCGGCCCGCGCGCGCCTCGCTCCCTCCCCGGGAACTCCGCTCTTGCGC
GGTCCCGGGCCTAGCGGCCAGCCTCGGCGGGACCGCACGTGCGCGCCCTCCCGGGCGGGG
GAGTCCCCGGGCGAGCACGTGCCGCGGCGCGCGGGGGGCGGGGAGTCCCGCGGCCGCCTT
CATTGCTCCTGCTGCTGCTGCCGCCGCCGCGGCCGCCGCCGCTGCTTCCTTCCTCCTGCG
CCGTCCCCTGCTCGGCCAGGGCTGGGGCGGGGGTTGGGGGGGGGGACTCCGCGCGCGCAG
CGAGCTTGGCTGGGCTGGGCGCTCCCGCTCCCGGAGGCGCACACCTTGGGCCGTGGCGGT
GGCGGTGGCGGCGGTGGCAGCGGCAGGCTGGGAAACTTTGCCCCTTTGTGCGCTCCGCCC
CGGAGGCGGCGGGCAGGTCGGTGCAGCCGGTGGCTTTCCGGCTCCGCGGTGGGGGAGGGG
CAGTGCGTGGGGGGAGGGGCGGCTCGGGGACCGTGGCACGGCCATGGTGGGCGCAGCAGG

//+600 base pairs of intronic sequence//

CCCCATTTTCTCTCTTCCCAGAAGCAGGAAGAGATAACACGTGAAAGTTGTCCCTGAGCA
CCCAGTGATGCACTTCCTCTTTCTGTGTCTTTTTCCCCTCTCAGGCAGCGTAGCCTTCTC
CATCATCCTGACCCAGGGAGGGAGAAGTGTCAGGAGGAAACTCCTCACCCCAATCTGCCC
CACCACACACCCATTGGCTCACCTCATGCCTGAGCCCCGGGTGGGTGAGGGTGAAGAACC
CACCCAGCCAAGATGAAATCTTCCTTGGGGGCTGCCGCCGCCACTGTCTTGCCCCAGATC
CTGGGGCCCTGCAGGTGGGAGACTGGTCCTGGAAGGAACCCTGTGAGCCGGCTGCAGAGG
AGGTGAAGAGGGTTAAGAAGAGGCACCCACCCCATCCCATCCCAGAGGCGGAGCTCCTTG
CCTTCACCCTTGTAGACTTGTTTAGGAGAGAAGGACCTGTTTGCAGGCAGCAGGAGACTG
GACAGTAGTTTAACCTGGCACCTATGCCCCCTTCCCTGGCAACGGCGGCATCATGAGCCA
GTTTCAGGTGCCCTTGGCGGTCCAGCCGGACCTGTCAGGACTTTATGATTTCCCGCAGGG
CCAGGTGATGGTAGGGGGCTTCCAGGGGCCTGGGCTTCCTATGGCTGGGAGTGAGACCCA

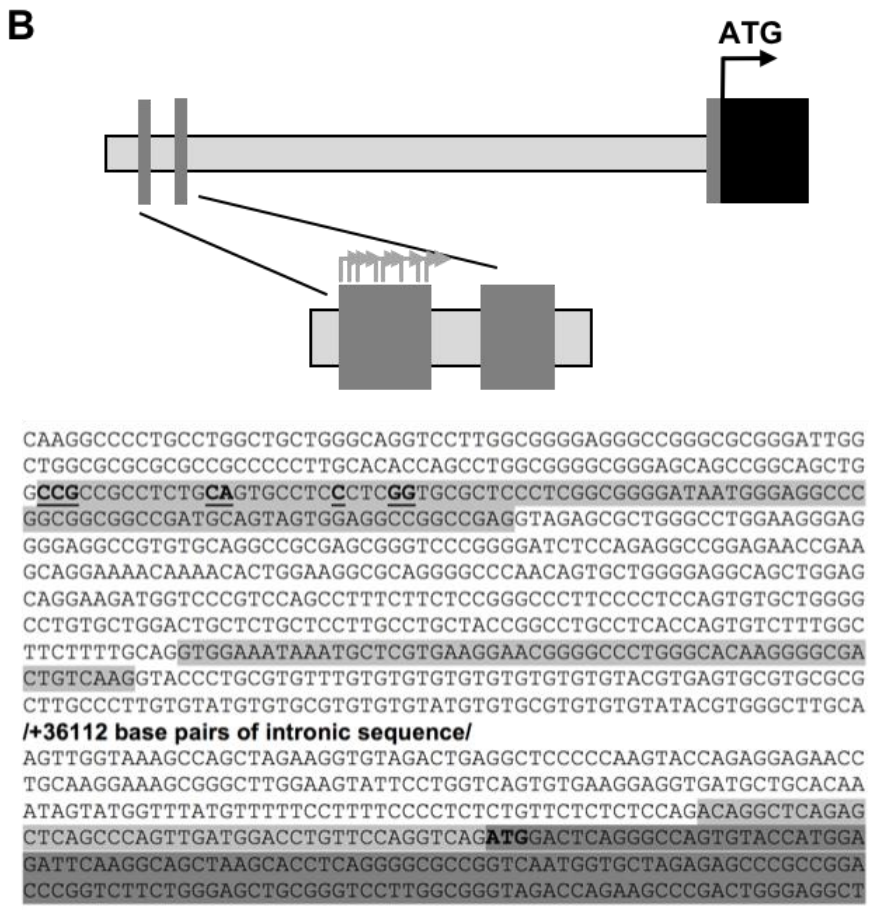


Figure 5:10 Identified cap sites for Tet3 upstream and downstream transcripts. 5'RACE was performed on day 7-differentiated mESCs to identify cap sites in 2 Tet3 protein isoforms (A) Tet3 upstream and (B) downstream. Identified cap sites are marked by arrows on the non-coding exons in the schematics and underlined in bold on the sequences. Light shading represents non-coding exons and darker shading represents the first coding exons (ATG codons are highlighted in bold).

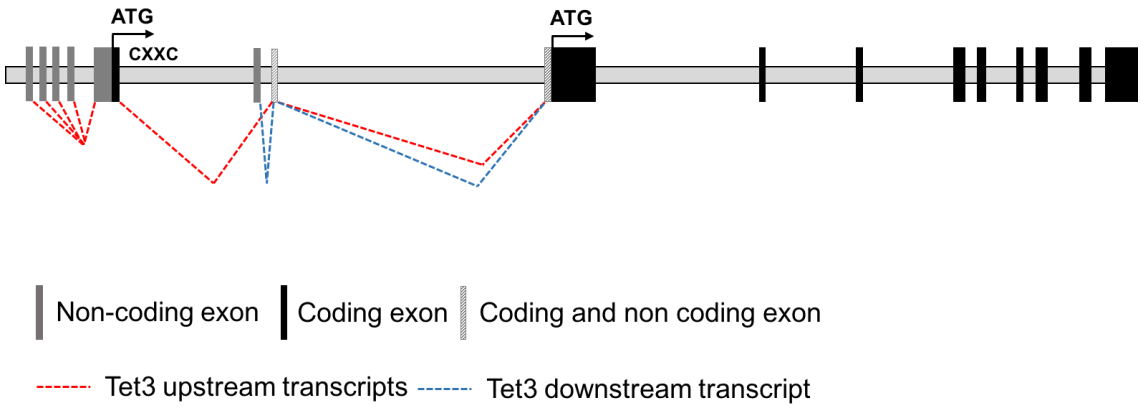


Figure 5:11 Tet3 has 2 distinct protein coding isoforms. 5'RACE, in combination with sequencing data available from (https://www.ensembl.org) and NCBI (https://www.ncbi.nlm.nih.gov), identified and confirmed multiple cap sites positioned upstream of 2 start codons. 4 non-coding exons were associated with the more upstream start site, compared to 1 non-coding exon associated with the downstream start. These transcripts marked as Tet3 upstream (red) and downstream (blue) give rise to 2 protein isoforms that differ by the presence or absence of a CXXC binding domain respectively.

5.3.5 mRNA expression of the Tet3 downstream transcript is induced in 1-week old mouse pup brain compared to heart

To quantitate the levels of the Tet3 transcripts that may arise from the 2 distinct promoter regions identified from 5'RACE, shown in section 5.3.4, primer pairs were designed for use in QPCR, to detect and distinguish between both the Tet3 upstream and Tet3 downstream variants (Figure 5:12).

Previously, the expression of all Tet3 transcripts (Tet3 pan) was shown to be induced in 1-week old brains compared to 1-week old hearts (Figure 5:6). The relative contributions of Tet3 upstream or downstream variants in mouse brain (and heart) were assessed by QPCR analysis (Figure 5:13). Tet3 upstream expression was lower, whereas Tet3 downstream expression was higher, in brain, compared to heart. Thus collectively, the elevated Tet3 pan expression in brain, compared to heart, is predominantly due to the downstream variant, suggestive that activation of the downstream promoter may be important in brain function.

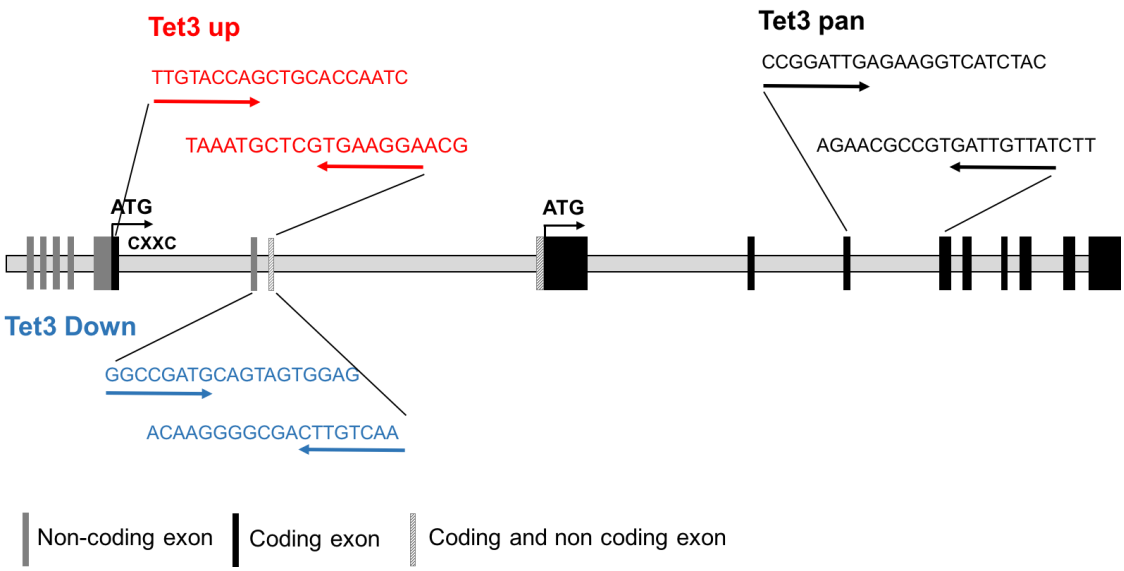


Figure 5:12 QPCR primer design to detect Tet3 transcript variants. Primers to detect specifically the Tet3 upstream and downstream variants were designed for QPCR analysis. Tet3 pan primers are also depicted to highlight their ability to detect both upstream and downstream variants.

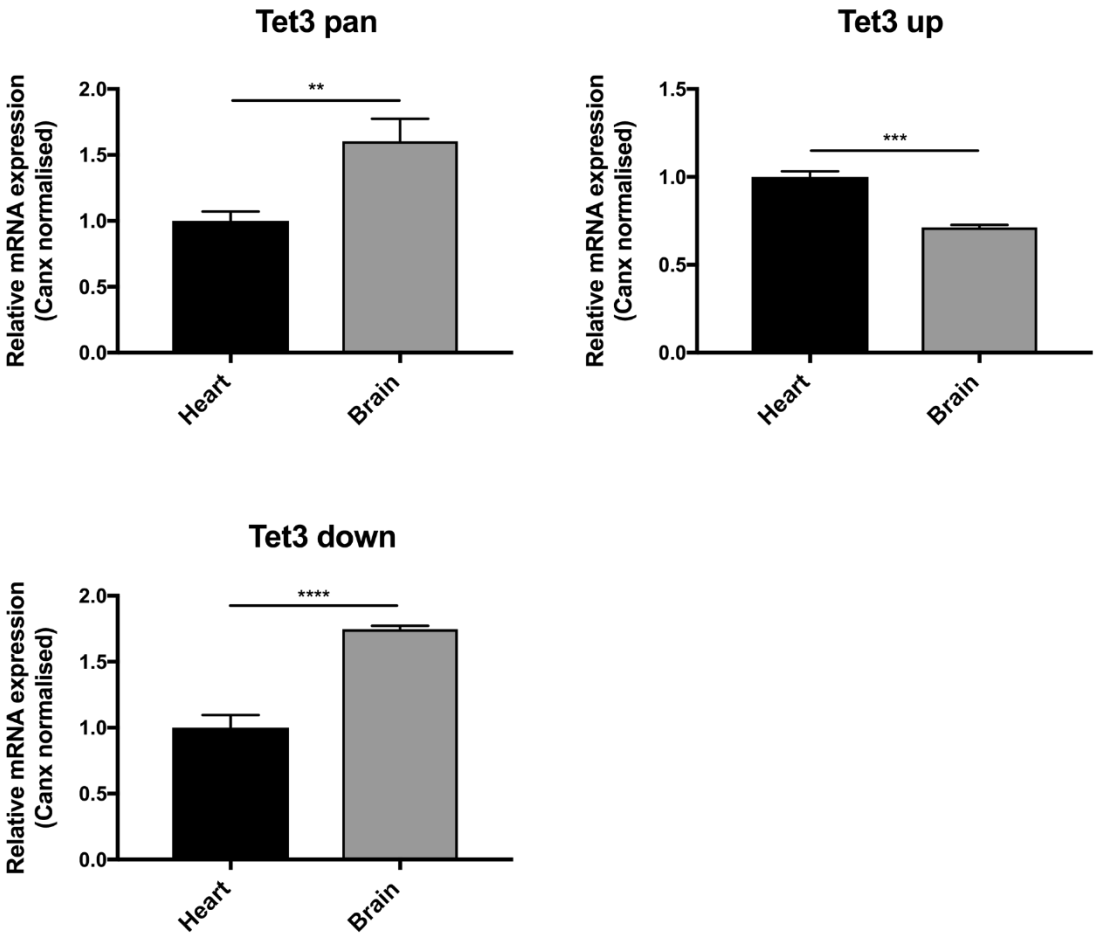


Figure 5:13 mRNA expression of the Tet3 downstream variant is elevated in 1-week old mouse pup brain, compared to heart. Relative mRNA expression of Tet3 pan, Tet3 upstream and Tet3 downstream in 1-week old mouse pup heart and brain was detected by QPCR analysis. Data are expressed as the mean ± SEM relative to heart expression, and analysed by an unpaired t-test. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$. n=3.

5.3.6 Elevated mRNA expression of the Tet3 downstream transcript appears partially sustained over a time course of neural-directed P19 cell differentiation, compared to the Tet3 upstream variant

To investigate the prospective involvement of the Tet3 downstream variant in neural specification, assessment of both upstream and downstream mRNA variants over a time course of neural and cardiac-directed P19 cell differentiation was established (Figure 5:14). Both upstream and downstream variants were induced in RA-, compared to DMSO-treated P19 cells. mRNA expression levels of both variants peaked at day 6, consistent with Tet3 pan expression (Figure 5:9). However, Tet3 downstream expression appeared to be sustained above mRNA levels detected in DMSO treated cells at days 8 and 10. Therefore, this may suggest the longer upstream protein isoform has a more transient role in neural-directed cellular differentiation.

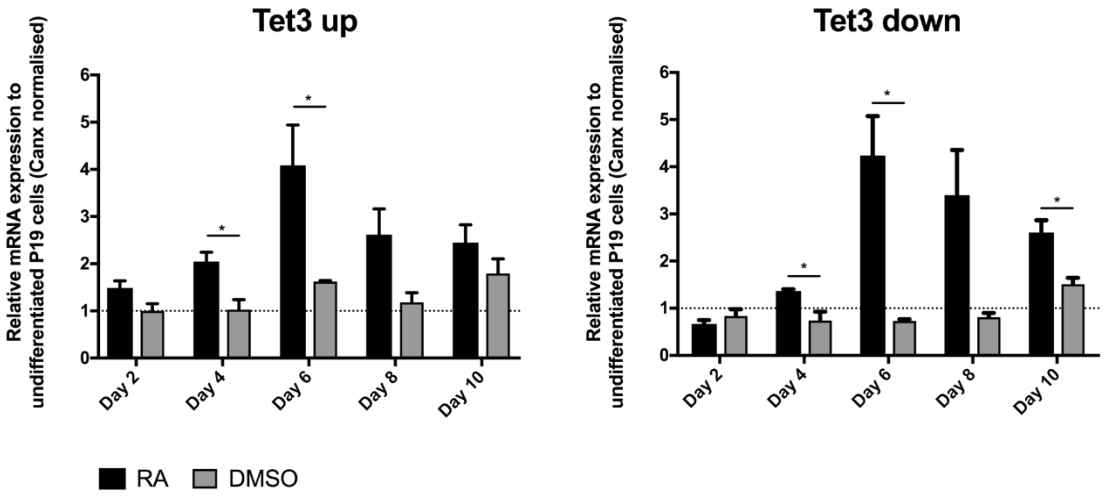


Figure 5:14 Both the Tet3 upstream and downstream variants are elevated in neural, compared to cardiac-directed P19 cell differentiation. P19 cells were induced to differentiate into neural and cardiac-like cells in the presence of 1 μM RA and 1% DMSO respectively. Expression of the Tet3 upstream and downstream variants were detected by QPCR. Data are expressed as the mean ± SEM relative to undifferentiated P19 cells, and analysed by multiple t-tests between RA and DMSO treated cells at each time point. * $p<0.05$

5.3.7 Tet3 upstream and Tet3 downstream mRNA are increased during mESC differentiation

To examine if either the Tet3 upstream or downstream variants show differential expression during mESC differentiation, thereby alluding to potential independent functions in cell specification events, their relative contributions to Tet3 pan expression were assessed by QPCR (Figure 5:15). Both Tet3 upstream and downstream correlated with the induction of Tet3 pan after day 3 of differentiation and seemed to both contribute to overall pan expression. The downstream variant appeared to be more highly expressed than the upstream in undifferentiated and day 3-differentiated mESCs, but at days 7 and 11 appeared equivalent. This therefore demonstrated that the upstream variant is induced to a greater extent than the downstream during this early period of cellular differentiation (this is also depicted in Figure 4:31).

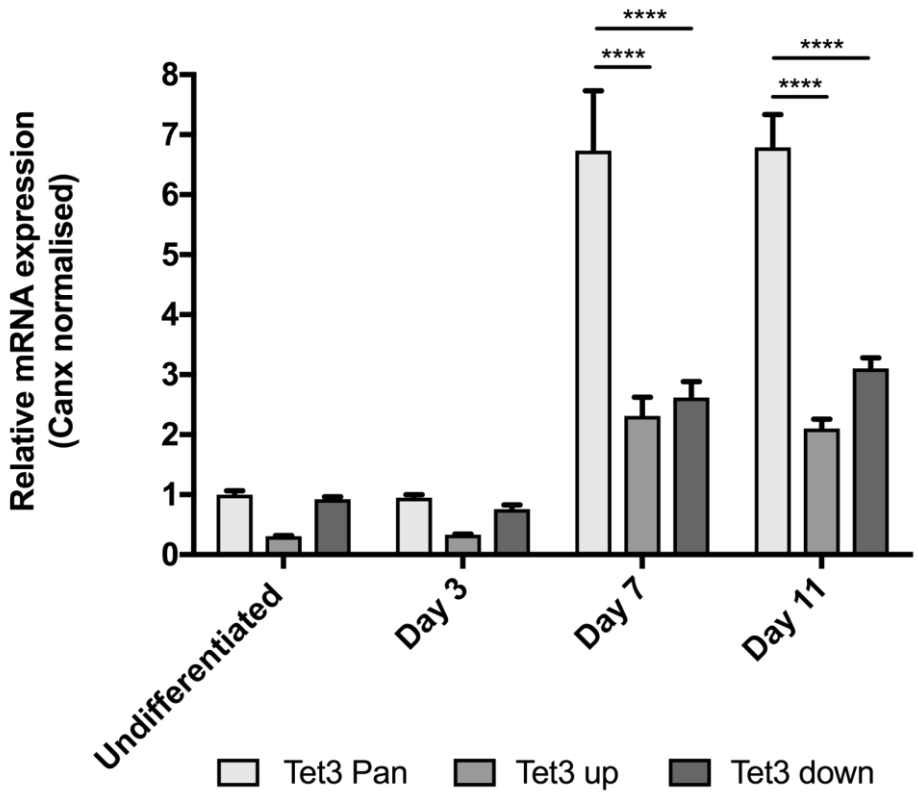


Figure 5:15 mRNA levels of Tet3 upstream and Tet3 downstream transcripts correlates with the induction of Tet3 pan expression after day 3 of mESC differentiation. The relative contributions of Tet3 upstream and Tet3 downstream to Tet3 pan expression over a time course of mESC differentiation was assessed by QPCR analysis. Data is expressed as the mean ± SEM relative to Tet3 pan in undifferentiated mESCs, and analysed by a two-way ANOVA with Tukey's post-hoc analysis. **** $p<0.0001$

5.3.8 Expression of the Tet3 downstream transcript is less sensitive to 1% O_2 than the Tet3 upstream variant

The loss of Tet3 pan mRNA expression in day 7 and 11-differentiated mESCs exposed to 1% O_2 (Figure 4:28) is consistent with the loss of neural makers under the same conditions (Figure 3:11, Figure 3:15 and Figure 3:18). The ability of the upstream and downstream transcripts to demonstrate differential sensitivities to 1% O_2 was examined by QPCR (Figure 5:16). Similarly to Tet3 pan expression, both transcripts were reduced by 1% O_2. However, low [O_2] appeared to have a less inhibitory effect upon the downstream variant. Examining expression differences between atmospheric O_2 and 1% O_2 revealed a 1.7 and 1.3-fold reduction at days 7 and 11 respectively for the downstream variant. In comparison, a 3- and 4-fold reduction at days 7 and 11 respectively was observed for the upstream variant.

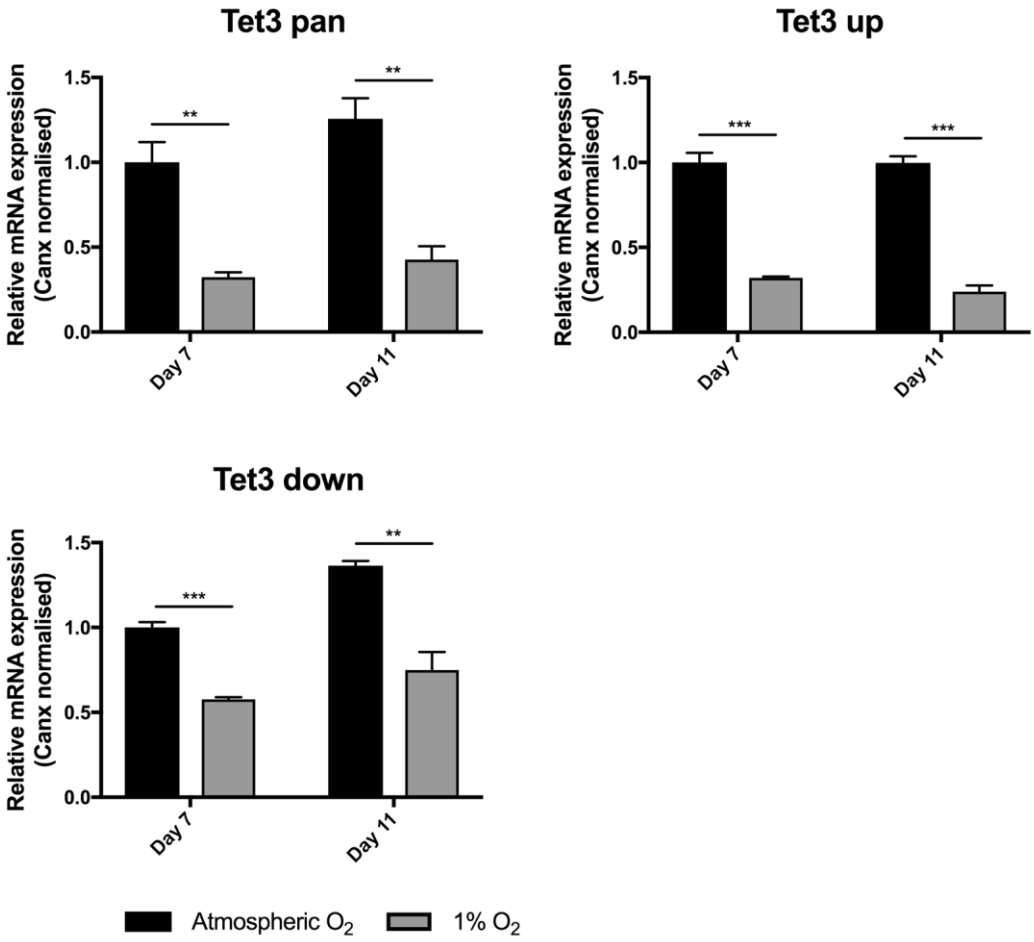


Figure 5:16 mRNA levels of the Tet3 downstream transcript is less sensitive to 1% O_2 than the Tet3 upstream variant. The sensitivities of Tet3 pan, Tet3 upstream and Tet3 downstream mRNA expression to 1% O_2 was determined in day 7 and 11-differentiated mESCs. Data are expressed as the mean ± SEM relative to day 7 atmospheric O_2 samples, and analysed by multiple t-tests between atmospheric O_2 and 1% O_2. ** $p<0.01$, *** $p<0.001$

5.3.9 The cloned Tet3 downstream putative promoter region has higher reporter activity, compared to the cloned Tet3 upstream promoter fragment, in both undifferentiated and differentiating P19 cells and mESCs

To investigate the transcriptional regulation of the 2 identified Tet3 transcripts, putative promoter regions were cloned into luciferase reporter constructs, schematically shown in Figure 5:17. The activity of these constructs were assessed in undifferentiated and day 2- and 4-differentiated P19 cells (Figure 5:18) and mESCs (Figure 5:19). Both putative promoter constructs were deemed functional, shown by induction of activity above the pGL4.2 control vector. In both cell types, the fragment of the cloned downstream promoter region consistently displayed higher levels of activity, compared to the upstream cloned promoter fragment, in both undifferentiated and differentiated cells. This may be consistent with the higher level of mRNA expression of the downstream, compared to the upstream transcripts in early differentiating mESCs (Figure 5:15). The promoter activity at later stages of cellular differentiation, where Tet3 expression was previously found to be induced (Figure 5:15), could not be investigated due to loss of the transient reporter expression after day 4. It can also be noted that the luciferase signals in mESCs were lower than those detected in P19 cells, likely owing to a lower transfection efficiency.

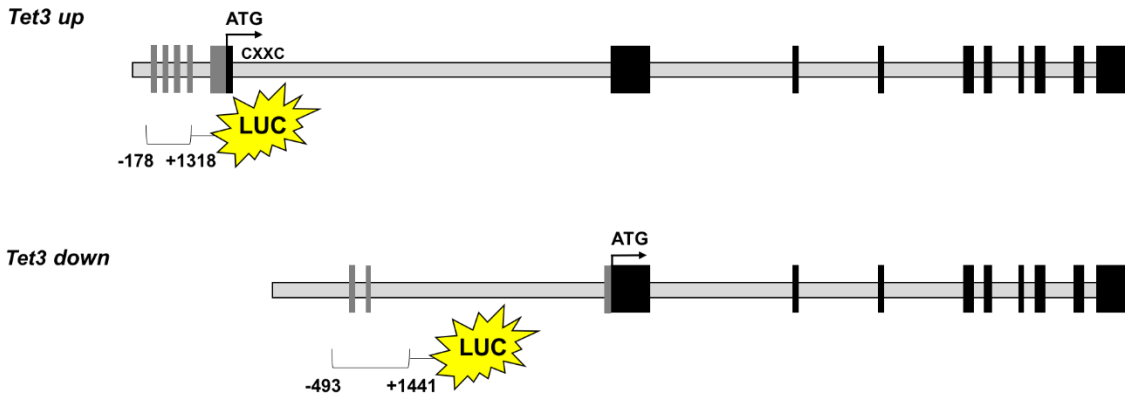


Figure 5:17 Schematic representation of the cloned putative promoter regions of the Tet3 upstream and downstream transcripts. The putative promoter regions of Tet3 upstream and Tet3 downstream transcripts were cloned into luciferase reporter constructs, as depicted. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5.10.

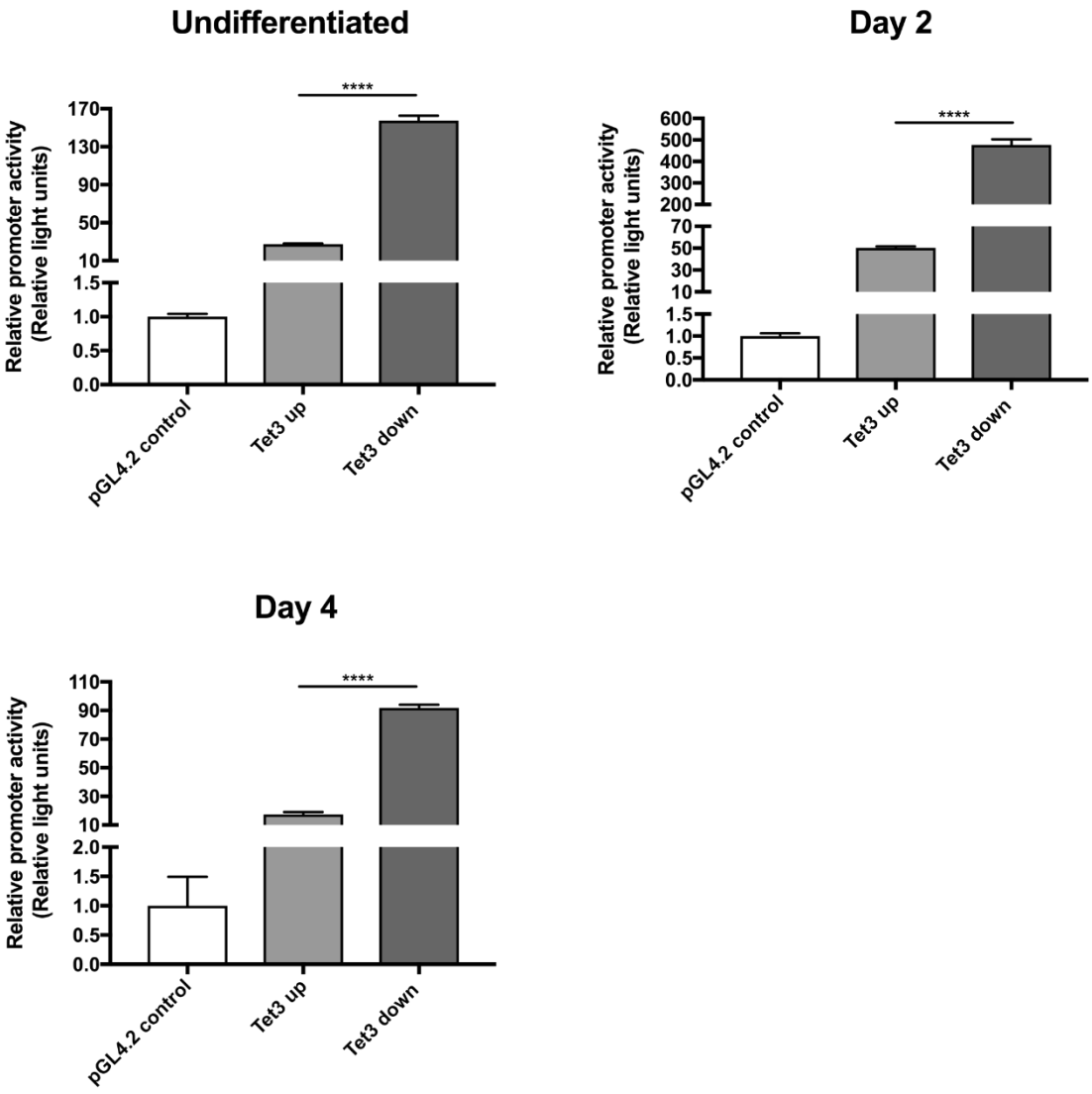


Figure 5:18 The Tet3 downstream and Tet3 upstream cloned putative promoter fragments showed elevated activity above the pGL4.2 control vector, in undifferentiated, day 2- and day 4-differentiated P19 cells. Luciferase reporter constructs containing the putative promoter regions of the Tet3 upstream and downstream transcripts were transfected into undifferentiated P19 cells. Luciferase activity was determined 24 hrs post-transfection or alternatively cells were differentiated and activity measured after 2 and 4 days. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. **** $p<0.0001$

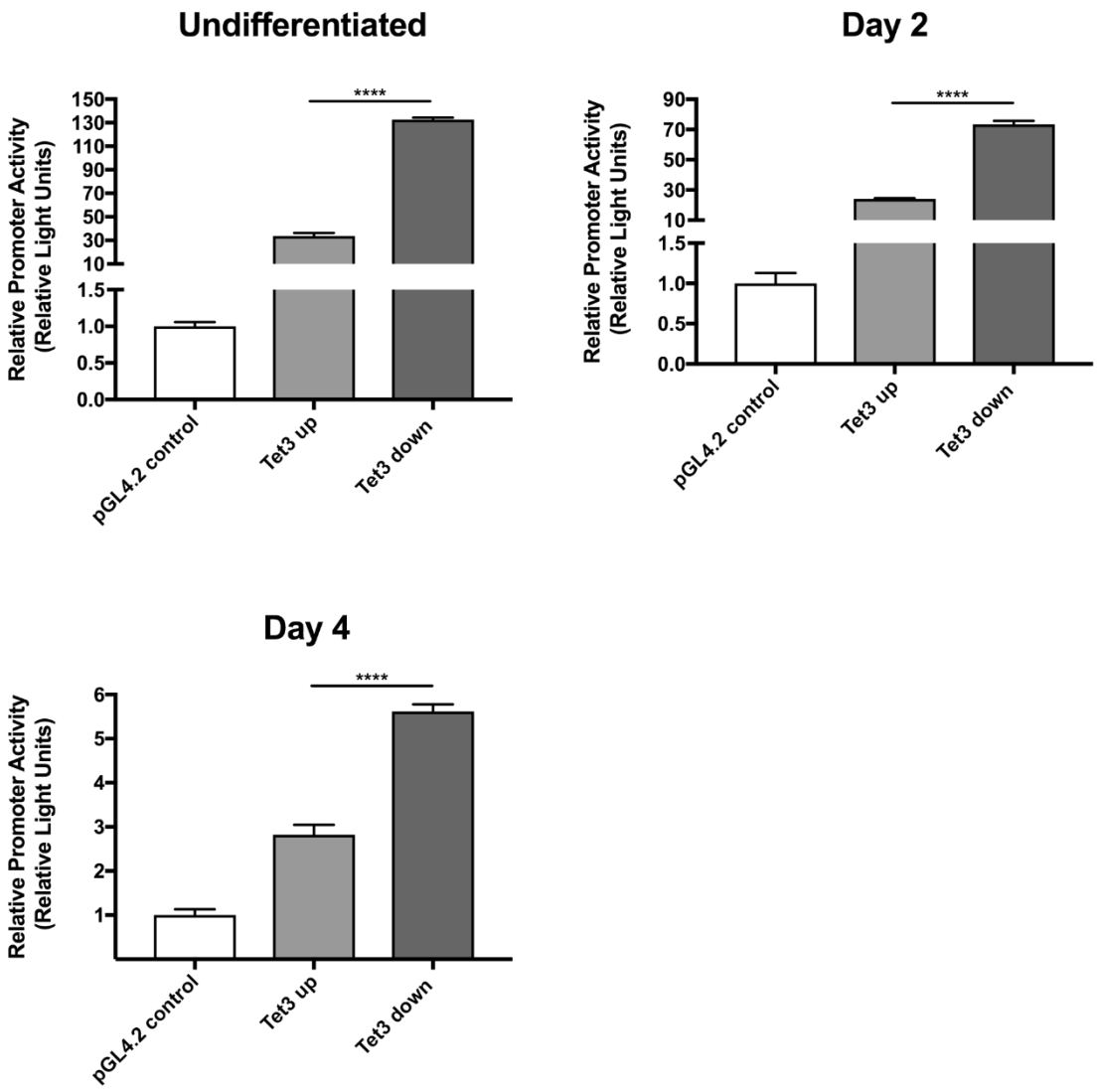


Figure 5:19 The Tet3 downstream and Tet3 upstream cloned putative promoter fragments showed elevated activity above the pGL4.2 control vector, in undifferentiated, day 2- and day 4-differentiated mESCs. Luciferase reporter constructs containing the putative promoter regions of the Tet3 upstream and downstream transcripts were transfected into undifferentiated mESCs. Luciferase activity was determined 24 hrs post-transfection or alternatively cells were differentiated and activity measured after 2 and 4 days. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. **** $p<0.0001$

5.3.10 A larger difference in reporter activity is evident between the Tet3 downstream and Tet3 upstream putative promoter region in neural, compared to cardiac cells

The activities of reporter constructs were assessed in more ‘differentiated’ cell types to gain further understanding of the mechanisms underlying the cell type specific expression patterns of the identified Tet3 variants. Tet3 upstream and downstream promoter reporter constructs were transfected into cardiac (H9c2) and neural (SH-SY5Y) cells types and luciferase activity determined (Figure 5:20). The downstream was more active in both tested cell types, and this was particularly evident in the neural cell type. Assessment of the ratio between the activities of the downstream and upstream promoter reporters in both cell types was determined from experimental repeats (Figure 5:20). The downstream promoter demonstrated an approximate 3-fold induction in SH-SY5Y cells, compared to an approximate 1.5-fold induction in H9c2 cells, above the upstream reporter. These data therefore are largely in agreement with elevated Tet3 downstream mRNA expression in brain, compared to heart tissue (Figure 5:13).

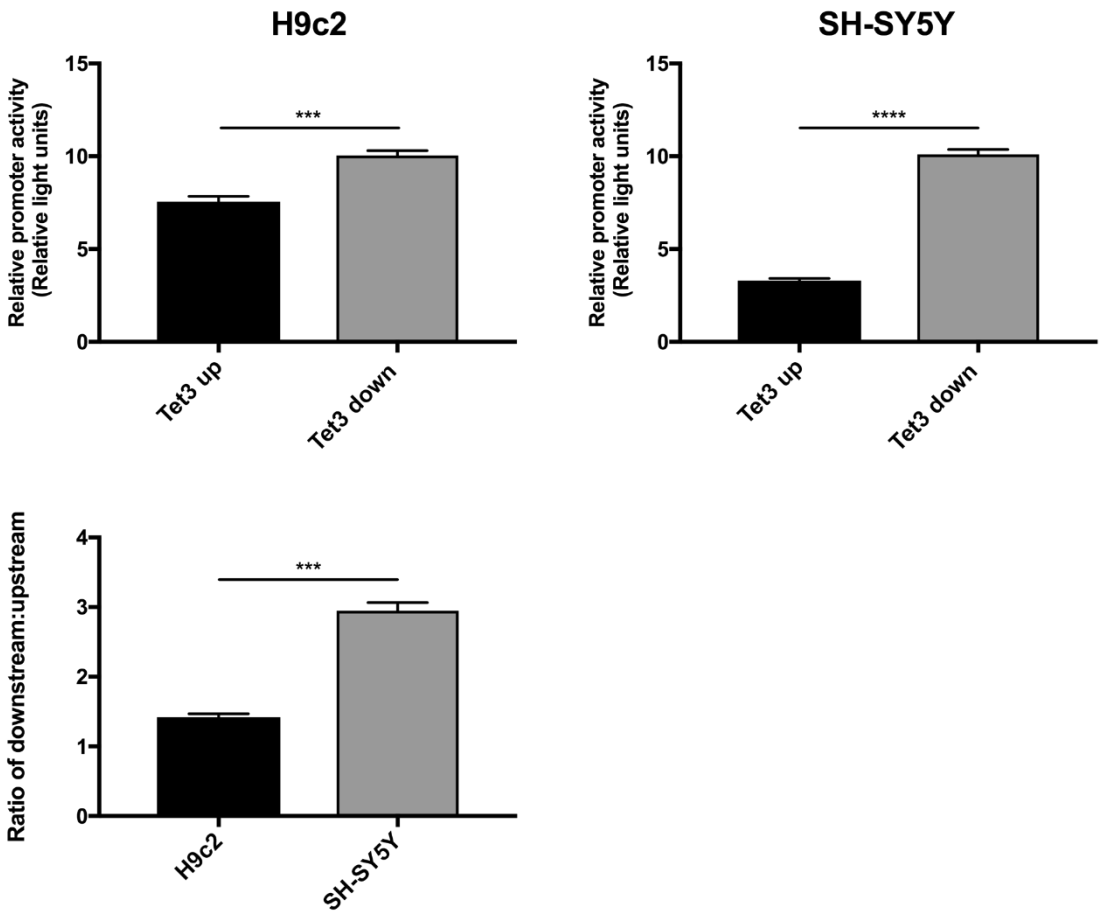


Figure 5:20 The activity of the Tet3 downstream putative promoter is elevated above the Tet3 upstream promoter to a greater extent in neural, compared to cardiac cells. Luciferase reporter constructs containing the putative promoter regions of Tet3 upstream and Tet3 downstream transcripts were transfected into H9c2 and SH-SY5Y cells, representative of cardiac and neural-like cells. Luciferase activity was determined 24 hrs post-transfection. The ratio of activity from the Tet3 downstream to the upstream reporter in H9c2 and SH-SY5Y cells, was calculated from 3 independent experiments. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by an unpaired t-test. *** $p<0.001$, **** $p<0.0001$

5.3.11 The Tet3 downstream putative promoter reporter construct is more active in 1% O_2, compared to atmospheric O_2, in undifferentiated P19 cells

The mRNA levels of both transcript variants of Tet3 were reduced in day 7- and 11-differentiated mESCs when exposed to 1% O_2, however expression of the downstream, compared to the upstream variant, appeared less sensitive to changes in [O_2] (Figure 5:16). To begin to assess whether the putative promoter activity of these transcript variants is also regulated by low [O_2], luciferase constructs were transfected into undifferentiated P19 cells and subjected to 1% O_2 (Figure 5:21). Perhaps surprisingly the Tet3 downstream promoter was shown to be elevated in 1% O_2, whereas the Tet3 upstream promoter appeared unaffected, when compared to atmospheric [O_2]. Assessment of the ratio between the downstream and upstream reporters in both tested O_2 tensions was determined from experimental repeats (Figure 5:21). The downstream promoter construct displayed an approximate 10-fold induction in 1% O_2 compared to approximately 6-fold in atmospheric O_2, above the upstream promoter. Interestingly, analysis of the downstream sequence revealed the presence of 8 core HIF DNA binding motifs (A/GCGTG), compared to 3 in the Tet3 upstream sequence. However, it should be noted this was conducted in undifferentiated P19 cells, it was not possible to look at the effects of [O_2] at later differentiated time points due to the loss of luciferase signal from the transiently transfected cells.

Attempts were made to investigate whether these effects were consistent in mESCs. However, owing to a low transfection efficiency, the pGL4.2 signal in 1% O_2 exposed cells was almost equivalent to the signal produced in untransfected cells (baseline) and thus these data are not shown.

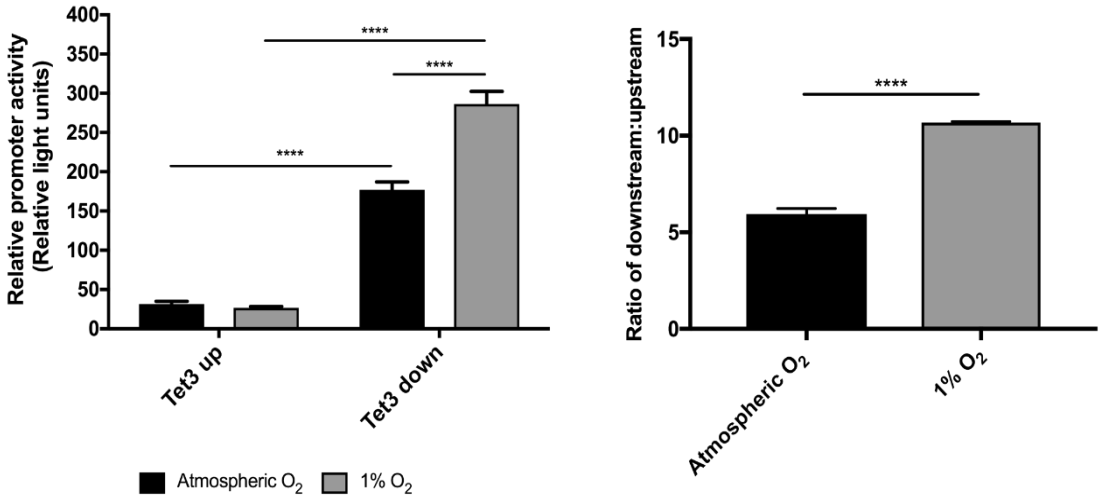


Figure 5:21 The activity of the Tet3 downstream putative promoter, compared to the Tet3 upstream, is induced in 1% O_2 compared to atmospheric O_2. Luciferase reporter constructs containing the putative promoter regions of Tet3 upstream and downstream transcripts were transfected into P19 cells and exposed to atmospheric O_2 or 1% O_2 for 24 hrs. The ratio of activity from the Tet3 downstream to the upstream reporter in atmospheric O_2 and 1% O_2 was calculated from 3 independent experiments. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a two-way ANOVA with Tukey's post-hoc analysis and an unpaired t-test. **** $p<0.0001$

5.3.12 A 5' and 3' deletion series of the Tet3 downstream putative promoter region decreased reporter activity

To identify the specific functional *cis*-acting promoter sequences that underlies potential neural specificity (shown in SH-SY5Y cells, see Figure 5:20) and O_2-dependent effects (Figure 5:21), a 5' and 3' deletion series of the Tet3 downstream reporter construct was conducted, schematically shown in Figure 5:22.

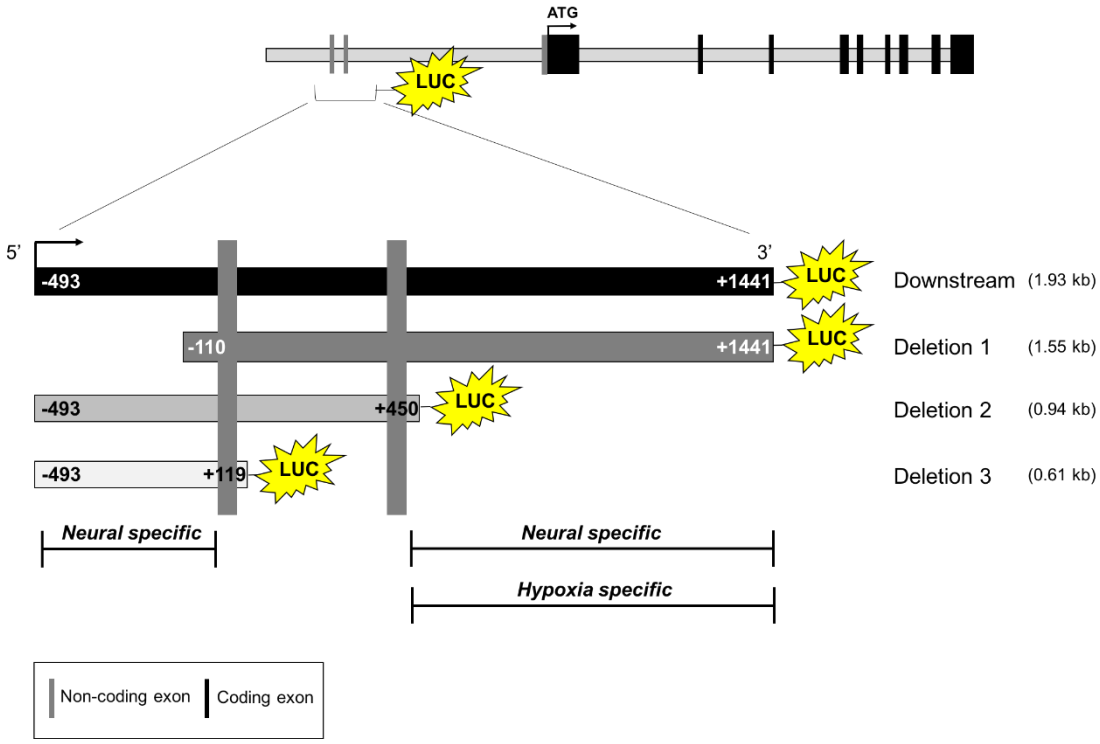


Figure 5:22 Schematic representation of the Tet3 downstream promoter deletion series. A deletion series of the putative Tet3 downstream promoter was conducted by sequence deletions from the 5' and 3' end. Specific regions were identified to have neural and O_2-dependent regulation as shown in Figure 5:25 and Figure 5:26. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5:10.

The effect of the deletion series in undifferentiated and day 2- and 4- differentiated P19 cells was assessed (Figure 5:23). In undifferentiated cells each deletion induced a reduction in promoter activity compared to the downstream construct, suggesting each deleted promoter sequence has a positive acting regulatory sequence. This effect was consistent in day 2 and day 4 differentiating cells, indicating there was not a distinct regulatory sequence required at the onset of differentiation. This is perhaps consistent with previous data in mESCs showing Tet3 expression does not increase until after day 3 of differentiation (Figure 5:15).

To confirm whether this deletion series was consistent in mESCs, the activity of the deletion constructs was assessed in undifferentiated cells (Figure 5:24). Consistent with undifferentiated P19 cells, each deletion had an inhibitory effect on promoter activity, compared to the downstream construct. Deletion 3, which is the shortest tested promoter fragment, produced the most pronounced reductive effect upon promoter activity.

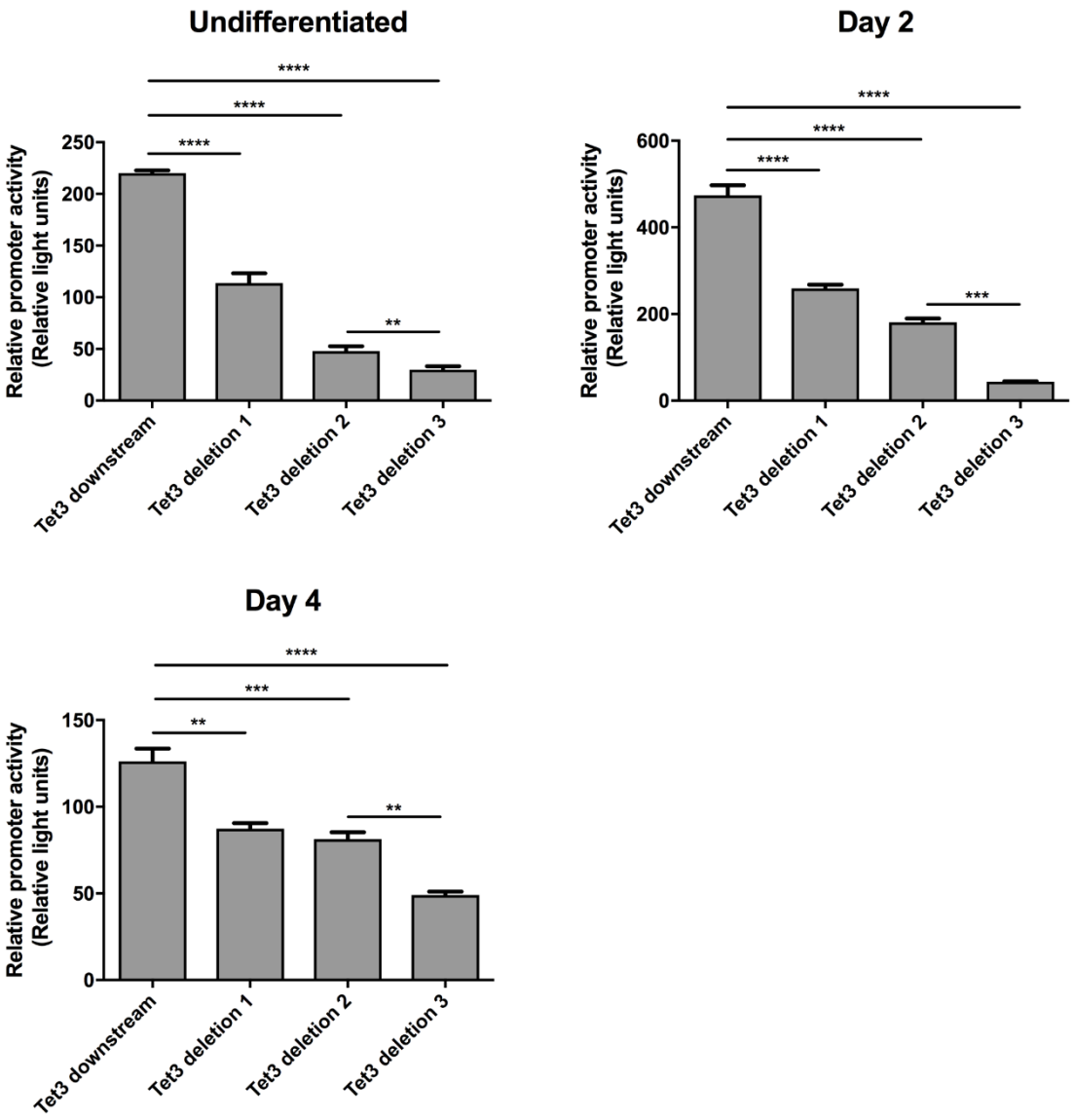


Figure 5:23 5' and 3' deletions from the Tet3 downstream promoter results in a decrease in reporter activity in undifferentiated and differentiated P19 cells. Luciferase reporter constructs containing the Tet3 downstream putative promoter or sequence deletions were transfected into undifferentiated P19 cells. Luciferase activity was determined 24 hrs post-transfection or alternatively cells were differentiated and activity measured after 2 and 4 days. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

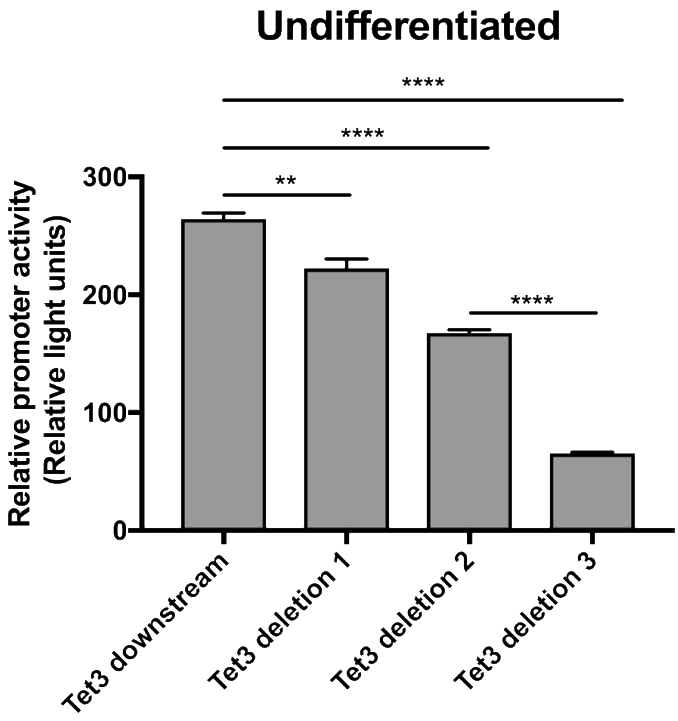


Figure 5:24 5' and 3' deletions from the Tet3 downstream promoter results in a decrease in reporter activity in undifferentiated mESCs. Luciferase reporter constructs containing the Tet3 downstream putative promoter or sequence deletions were transfected into undifferentiated mESCs. Luciferase activity was determined 24 hrs post-transfection. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

5.3.13 A deletion series of the Tet3 downstream putative promoter identified potential neural-specific regulatory fragments

The Tet3 downstream putative promoter fragment demonstrated a higher relative activity, compared to the upstream promoter region, in neural (SH-SY5Y) cells compared to cardiac (H9c2) cells (Figure 5:20). To identify cis acting sequence(s) that might potentially promote expression specifically in neural cells, the activity of the Tet3 downstream promoter deletion reporter constructs were assessed in both SH-SY5Y and H9c2 cells (Figure 5:25). All the tested deleted reporter constructs reduced activity in SH-SY5Y cells, compared to the undeleted construct, but increased (or maintained) activity in H9c2 cells. Specifically, deletion 1 and 2 induced an equivalent 50% reduction in activity in SH-SY5Y cells, compared to a significant induction in activity in H9c2 cells. Deletion 3, compared to Deletion 2, reduced activity in both cell types, thereby identifying

common positive acting sequence(s) within this deleted region. Overall, 2 potential cis-acting promoter sequences, specific for Tet3 promoter activity in a neural cell type was identified, marked in the schematic shown in Figure 5:22.

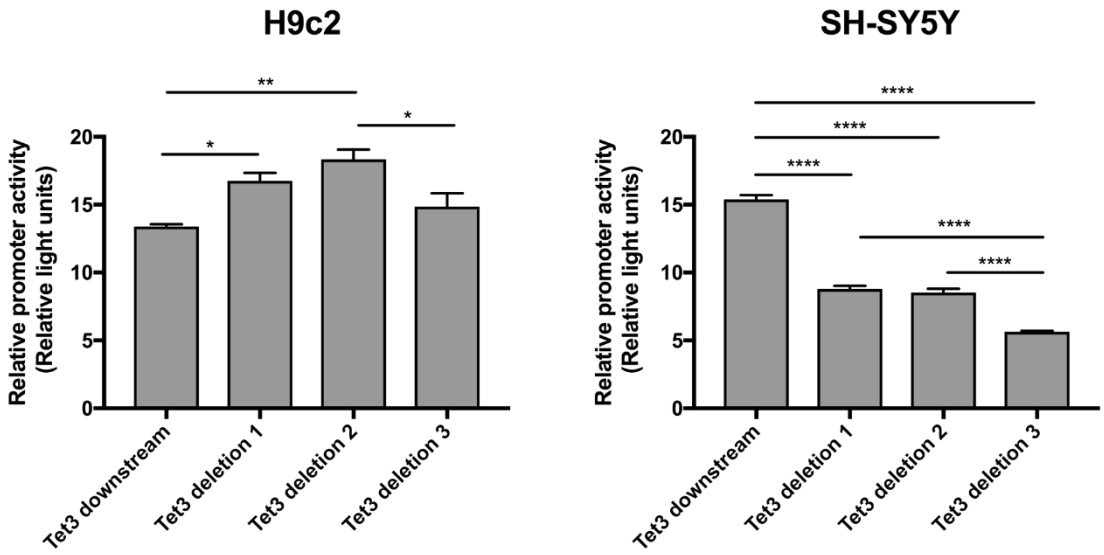


Figure 5:25 Deletions of the Tet3 downstream putative promoter identified 2 neural-specific regulatory fragments. Luciferase reporter constructs containing the Tet3 downstream putative promoter or sequence deletions were transfected into H9c2 and SH-SY5Y cells. Luciferase activity was determined 24 hrs post transfection. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. Shown are representative figures from 3 independent repeats. ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$

5.3.14 A deletion series of the Tet3 downstream putative promoter identified an O_2-dependent regulatory region

The Tet3 downstream putative promoter construct demonstrated an induction in activity when subjected to 1% O_2 in undifferentiated P19 cells (Figure 5:21). To determine the potential O_2-senstive fragment(s), the activity of Tet3 downstream deletion reporter constructs were assessed in P19 cells exposed to 1% O_2 (Figure 5:26). The activity of the Tet3 downstream construct was confirmed to be increased in 1% O_2, an effect maintained in deletion 1, yet lost in deletions 2 and 3. This therefore identified a potential O_2-dependent regulatory region, depicted on the schematic shown in Figure 5:22.

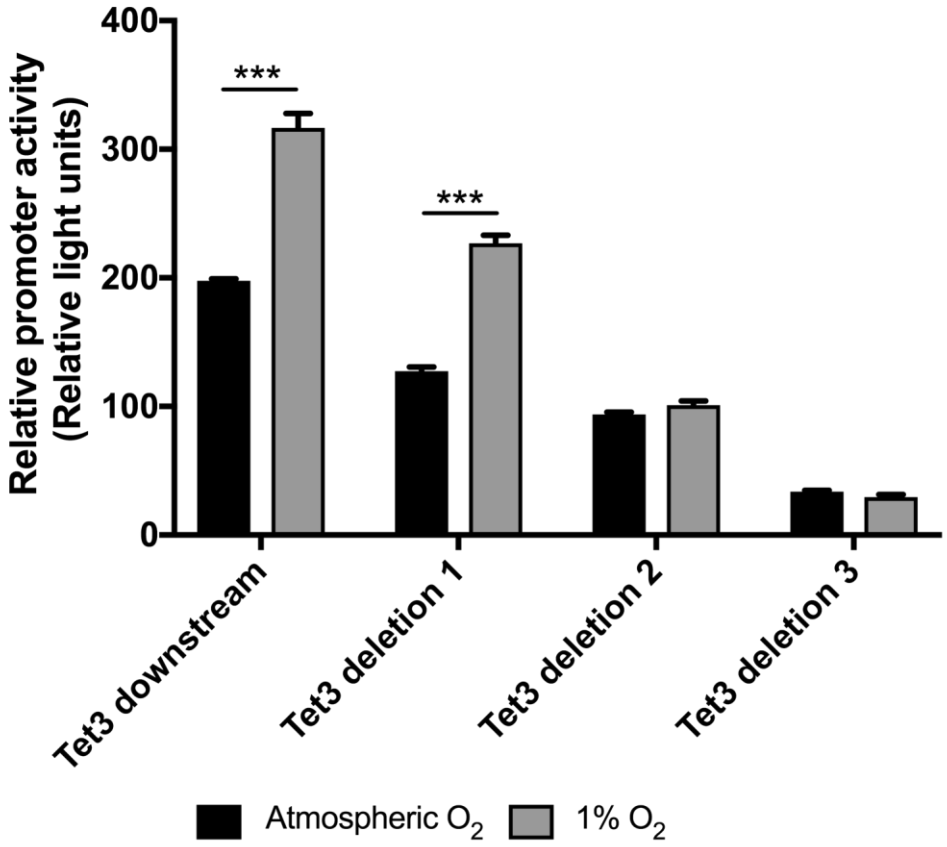


Figure 5:26 Deletions of the Tet3 downstream putative promoter identified an O_2–sensitive regulatory region. Luciferase reporter constructs containing the Tet3 downstream putative promoter or sequence deletions were transfected into undifferentiated P19 cells and exposed to atmospheric O_2 or 1% O_2 for 24 hrs. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by multiple t-tests between atmospheric O_2 and 1% O_2. *** $p<0.001$

5.3.15 A 3' extension of the Tet3 upstream putative promoter region has a positive regulatory function in undifferentiated P19 and mESCs

The Ensembl database (https://www.ensembl.org) indicates a putative promoter region that encompasses a large intronic region on the mouse Tet3 upstream variant. Thus, to characterise additional potential regulatory sequences, which are not present in the existing upstream promoter fragment, 5' and 3' sequence extensions, in addition to a deletion were cloned, shown schematically in Figure 5:27.

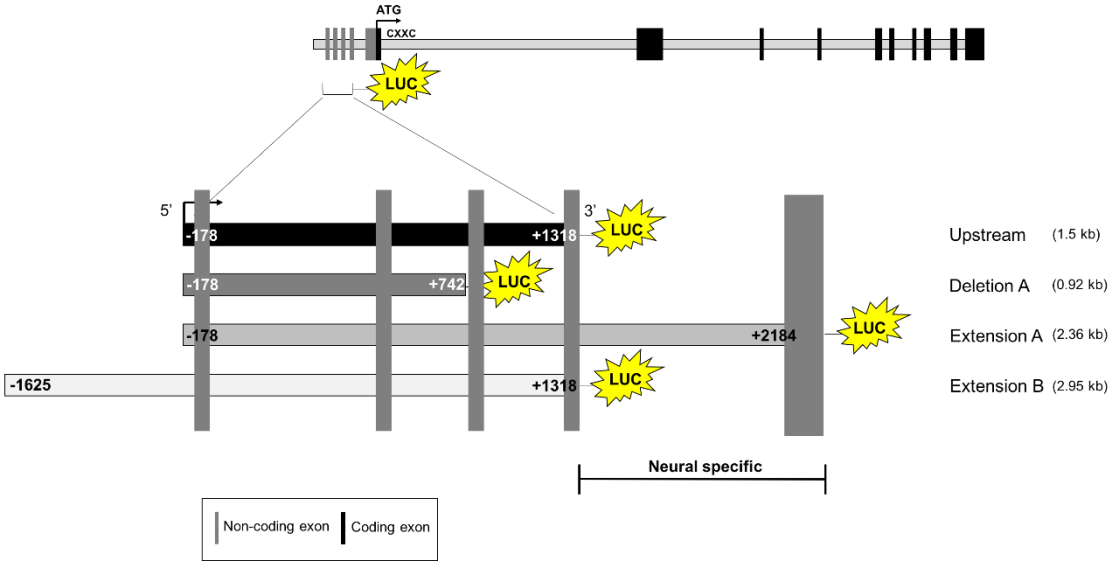


Figure 5:27 Schematic representation of 5' and 3' extensions, and a deletion, of the Tet3 upstream putative promoter. Extensions containing potential regulatory sequence at the 5' and 3' end of the Tet3 upstream putative promoter were cloned into luciferase reporters. A deletion construct was also created from 3' end of the Tet3 upstream putative promoter. A specific region identified to have neural-specificity, as shown in Figure 5:29, is marked up. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5.10.

The activity of these constructs was first assessed in undifferentiated P19 and mESCs (Figure 5:28). A 3' extension of the Tet3 upstream putative promoter sequence, named extension A, had a positive regulatory effect in both P19 and mESCs. However, it should be noted that in P19 cells the downstream promoter fragment remained more active than the upstream extension A fragment. In addition, the positive regulatory effect of the 3' extended sequence, relative to the Tet3 upstream promoter construct, was markedly greater in mESCs, compared to P19 cells (9.3-fold compared to 1.7-fold). Deletion A reduced promoter activity, compared to the Tet3 upstream promoter, consistently in both

cell types, suggesting this 3' end of the putative promoter (encompassing one of the identified cap site regions) contains a positive regulatory sequence. Lastly, the 5' extended genomic fragment, extension B, had a marginal inhibitory effect on activity in P19 cells and no effect in mESCs, compared to the original Tet3 upstream construct. Thus, there are cis-acting sequences downstream of the identified cap sites that act to promote transcription of Tet3 in undifferentiated P19 cells and mESCs.

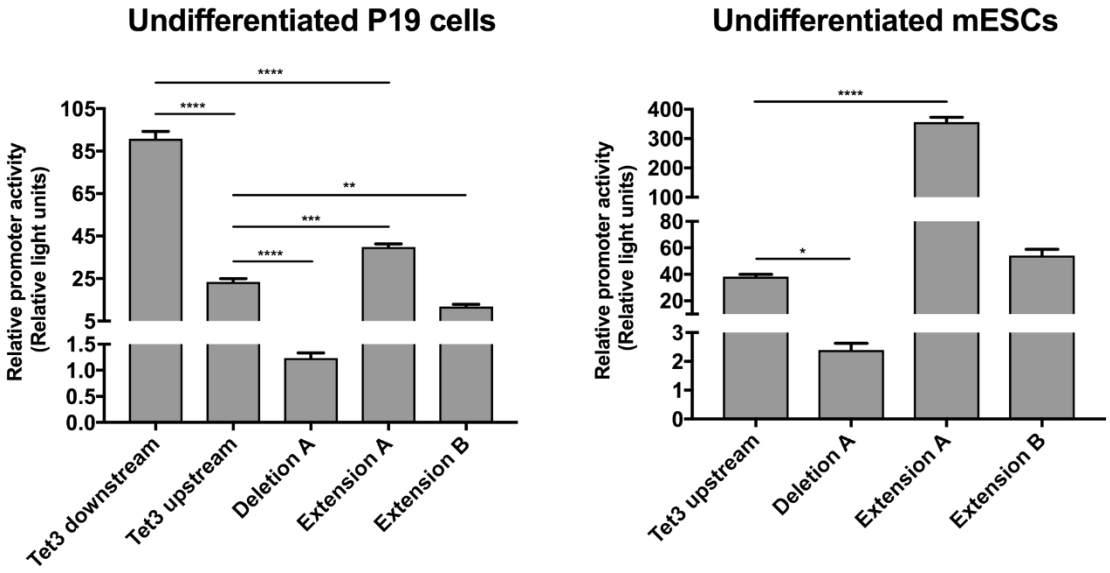


Figure 5:28 A 3' end extension of the Tet3 upstream promoter has a positive regulatory function in undifferentiated P19 cells and mESCs. Luciferase reporter constructs containing the Tet3 upstream putative promoter or the sequence extensions/deletion were transfected into undifferentiated P19 cell and mESCs. Note the Tet3 downstream promoter was also tested in undifferentiated P19 cells. Luciferase activity was determined 24 hrs post-transfection. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's (for P19 cells) or Dunnett's (for mESCs) post-hoc analysis. * $p<0.05$, ** $p<0.01$, *** $p<0.001$, **** $p<0.0001$.

5.3.16 The 3' extension of the Tet3 upstream promoter has a neural specific regulatory element

The mRNA level of the Tet3 upstream transcript showed increased levels in neural-directed, compared to cardiac-directed P19 cell differentiation (Figure 5:14). In addition, this transcript was shown to be induced during early mESC differentiation to greater extent than the downstream transcript (Figure 5:15). To establish whether the Tet3 upstream promoter reporter construct could function in a cell-type specific manner, the activities of the series of upstream reporter constructs was assessed in H9c2 and SH-SY5Y cells (Figure 5:29). Deletion A and extension B reduced luciferase activity, compared to the Tet3 upstream construct, in both cell types, identifying positively and negatively acting regulatory regions respectively. However, extension A induced an approximate 6-fold induction in activity in SH-SY5Y cells above the upstream construct. By contrast, in H9c2 cells extension A showed no difference in promoter activity compared to the upstream reporter. Therefore, it can be concluded that the sequence comprising this 3' extension of the Tet3 upstream putative promoter has a neural-specific regulatory function, marked on the schematic shown in Figure 5:29.

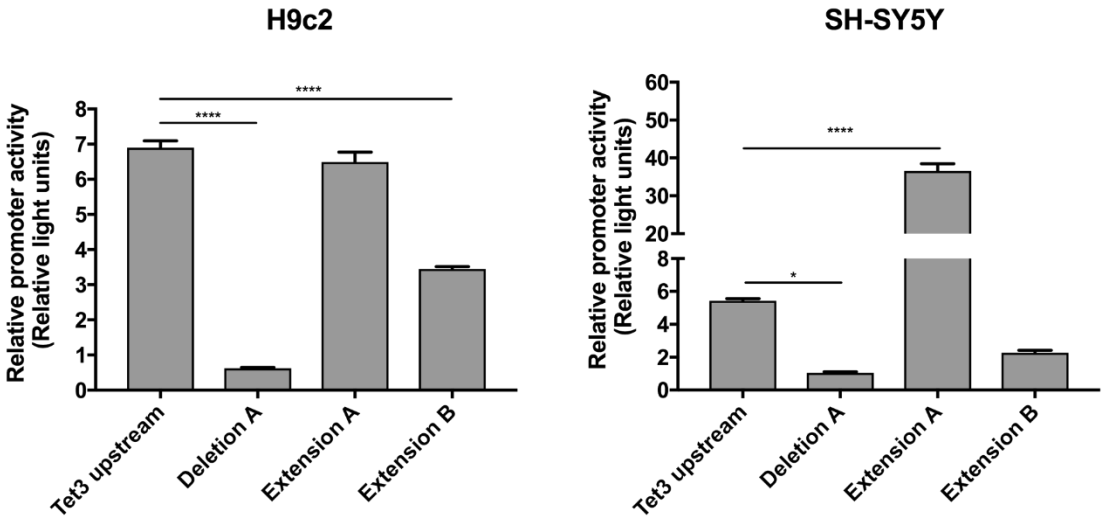


Figure 5:29 A 3' extension of the Tet3 upstream putative promoter has a neural-specific regulatory function. Luciferase reporter constructs containing the Tet3 upstream putative promoter or the sequence extensions/deletion were transfected into H9c2 and SH-SY5Y cells. Luciferase activity was determined 24 hrs post-transfection. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Dunnett's post-hoc analysis. Shown are representative figures from 3 independent repeats. * $p<0.05$, **** $p<0.0001$.

5.3.17 Deletions from the neural-specific 3' extension of the Tet3 upstream putative promoter identified a 160-base pair regulatory sequence

To identify the precise regulatory sequence responsible for the neural-specific effect of this 3' extension, 200 bp deletions were made from the extension A sequence (Figure 5:30).

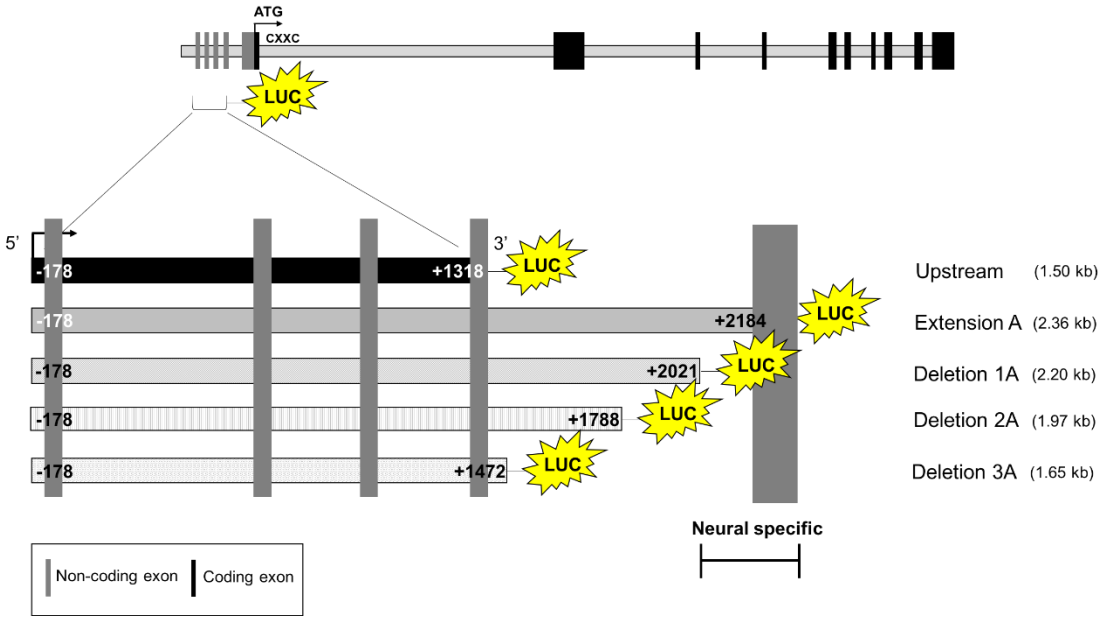


Figure 5:30 Schematic representation of extension A deletions. Consecutive 200 bp deletions at the 3' end of the extension A sequence were cloned into luciferase reporter constructs. An identified potential neural specific region is marked as shown in Figure 5:31. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5.10.

The promoter activity of this deletion series was assessed in H9c2 and SH-SY5Y cells (Figure 5:31). Deletion 1A reduced promoter activity, compared to extension A, by approximately 9-fold in SH-SY5Y cells and 5-fold in H9c2 cells. Deletion 2A increased promoter activity in both cell types, compared to deletion 1A (activity remained reduced compared to extension A), suggestive that this 230 bp region is inhibitory. A further deletion from extension A, deletion 3A, also had an inhibitory effect in both cell types. However, compared to deletion 1A, a marginal, but significant elevation in H9c2 cells, which was not seen in SH-SY5Y cells, was apparent. Therefore, the major regulatory sequence was identified to be found between extension A and deletion 1A. Despite also reducing luciferase activity in H9c2 cells, a larger fold reduction was detected in SH-

SY5Ys with deletion 1A, suggestive that this sequence may have some neural specific function (marked on the schematic in Figure 5:30).

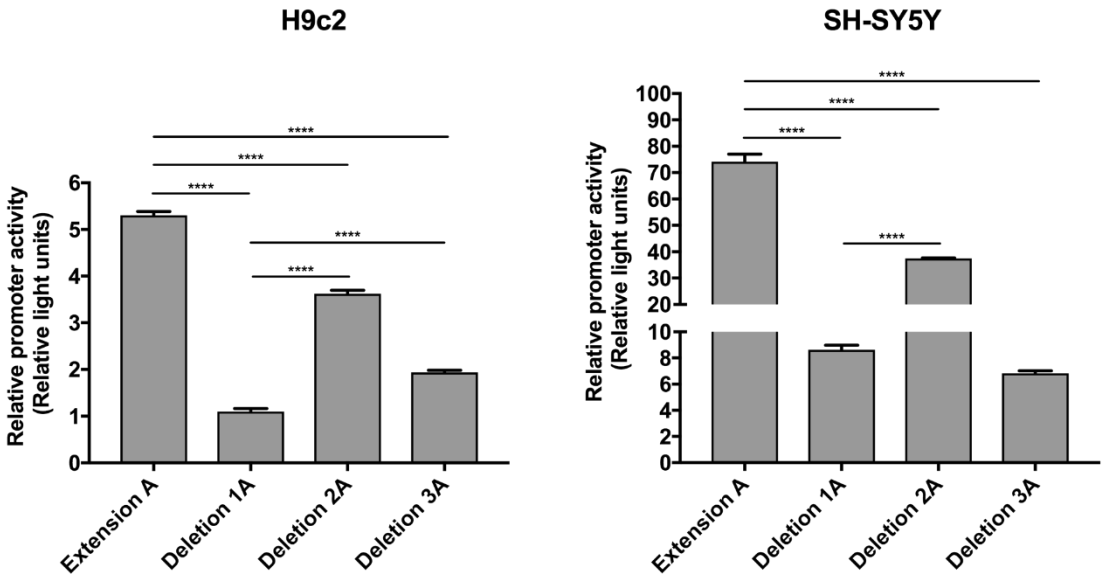


Figure 5:31 Deletion 1A reduces luciferase activity to a larger order of magnitude in neural cells than cardiac cells, compared to extension A. Luciferase reporter constructs containing extension A and its deletions were transfected into H9c2 and SH-SY5Y cells. Luciferase activity was determined 24 hrs post-transfection. Data are expressed as the mean ± SEM relative to the activity of pGL4.2 control vector after normalisation to the co-transfected Renilla luciferase control. Data are analysed by a one-way ANOVA with Tukey's post-hoc analysis. Shown are representative figures from 3 independent repeats. **** $p<0.0001$.

5.3.18 Sequence homology is found between the identified neural-specific regulatory promoter fragments in the mouse compared to human

Similar to the mouse, the human Tet3 gene also encodes variants with a CXXC domain (NCBI accession numbers: NM_001287491.1, XM_011532689.2, XM_011532688.2, XM_011532686.2, XM_011532685.2, XM_005264187.3, XM_011532684.2, XM_011532683.2, XM_017003566.1, XM_011532682.2 and XM_011532690.1) and without a CXXC domain (NCBI accession numbers: XM_017003567.1 and XM_011532687.2). Sequence comparison of the cloned mouse putative Tet3 promoter fragments (Tet3 upstream co-ordinates: -1625 to +2184 and Tet3 downstream co-ordinates: -493 to +1441) and the homologous human sequences, conducted using the Clustal Omega multiple sequence alignment tool (https://www.ebi.ac.uk/Tools/msa/clustalo/), revealed an estimated overall similarity of 70% and 80% for the upstream and downstream promoter regions respectively. Alignment of the identified murine neural-regulatory fragments in the Tet3 downstream and upstream putative promoters to the human sequence was also conducted (Figure 5:32 and Figure 5:33). The 2 neural-specific sequences for the downstream promoter, highlighted schematically in Figure 5:22, demonstrated the more 3' promoter fragment (Figure 5:32B) has a higher homology to the human sequence. The neural-specific sequence identified from deletions of the extension A of the upstream putative promoter fragment, schematically depicted in Figure 5:30, also appears to have a high homology to the human sequence (Figure 5:33). Therefore, this high level of homology between mouse and human sequences may suggest these identified neural-specific fragments are highly conserved and provides a translational basis to study Tet3 regulation further.

A

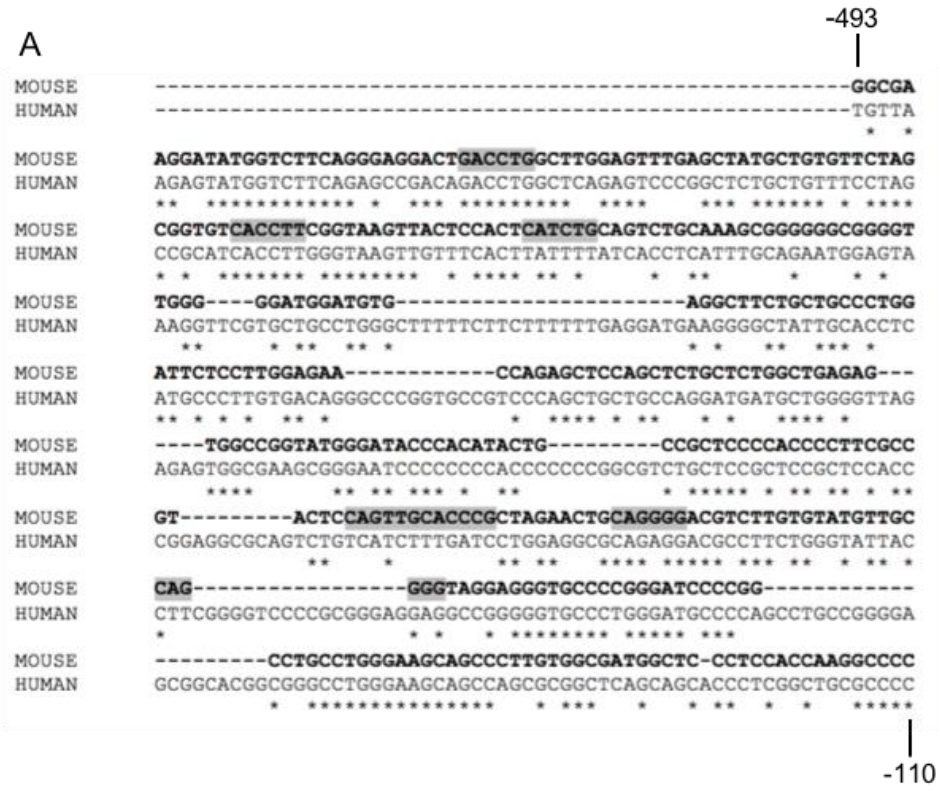

-493
MOUSE ----------GGCGA
HUMAN ----------TGTTA
MOUSE AGGATATGGTCTTCAGGGAGGACTGACCTGGCTTGGAGTTTGAGCTATGCTGTGTTCTAG
HUMAN AGAGTATGGTCTTCAGAGCCGACAGACCTGGCTCAGAGTCCCGGCTCTGCTGTTTCCTAG
MOUSE CGGTGTCACCTTCGGTAAGTTACTCCACTCATCTGCAGTCTGCAAAGCGGGGGGCGGGGT
HUMAN CCGCATCACCTTGGGTAAGTTGTTTCACTTATTTTATCACCTCATTTGCAGAATGGAGTA
MOUSE TGGG----GGATGGATGTG----------AGGCTTCTGCTGCCCTGG
HUMAN AAGGTTCGTGCTGCCTGGGCTTTTTCTTCTTTTTTGAGGATGAAGGGGCTATTGCACCTC
MOUSE ATTCTCCTTGGAGAA----------CCAGAGCTCCAGCTCTGCTCTGGCTGAGAG---
HUMAN ATGCCCTTGTGACAGGGCCCGGTGCCGTCCCAGCTGCTGCCAGGATGATGCTGGGGTTAG
MOUSE ----TGGCCGGTATGGGATACCCACATACTG---------CCGCTCCCCACCCCTTCGCC
HUMAN AGAGTGGCGAAGCGGGAATCCCCCCCCCACCCCCCCGGCGTCTGCTCCGCTCCGCTCCACC
MOUSE GT---------ACTCCAGTTGCACCCGCTAGAACTGCAGGGGACGTCTTGTGTATGTTGC
HUMAN CGGAGGCGCAGTCTGTCATCTTTGATCCTGGAGGCGCAGAGGACGCCTTCTGGGTATTAC
MOUSE CAG----------GGGTAGGAGGGTGCCCCGGGATCCCCGG----------
HUMAN CTTCGGGGTCCCCGCGGGAGGAGGCCGGGGGTGCCCTGGGATGCCCCAGCCTGCCGGGGA
MOUSE ---------CCTGCCTGGGAAGCAGCCCTTGTGGCGATGGCTC-CCTCCACCAAGGCCCC
HUMAN GCGGCACGGCGGGCCTGGGAAGCAGCCAGCGCGGCTCAGCAGCACCCTCGGCTGCGCCCC
-110

Figure 5:32 The 3' neural-specific fragment of Tet3 downstream putative promoter has high sequence homology with the human sequence. Alignment of the murine Tet3 downstream neural-specific putative promoter fragments with the human sequence are shown. (A) The neural region identified at the 5' end of the downstream construct (co-ordinates: -493 to -110). (B) The neural region identified at the 3' end of the downstream construct (co-ordinates: +450 to +1441). The mouse sequence is highlighted in bold. * indicates sequence homology. Grey shading and underlining highlights putative binding sites for the TF Snal1 and Sox5 respectively, see section 5.3.19. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5.10.

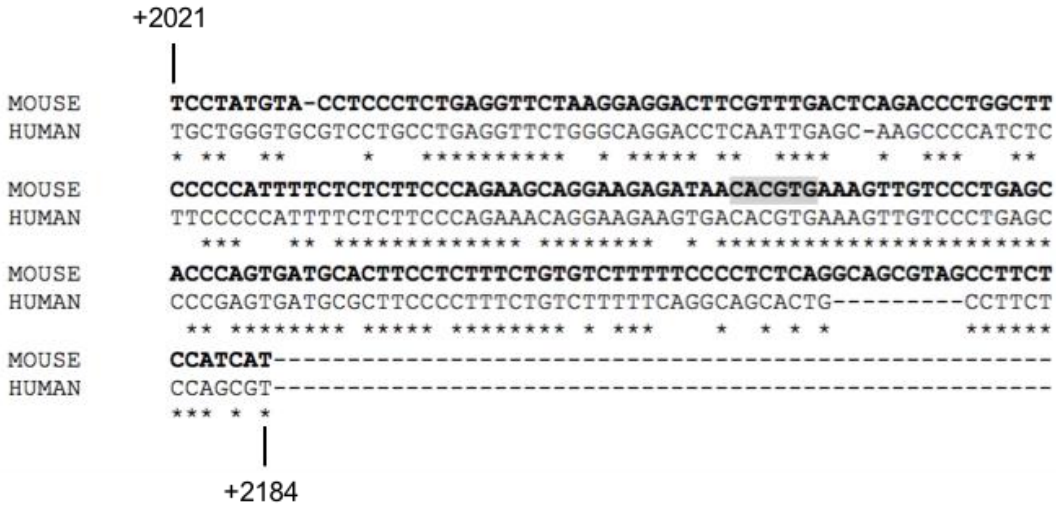


Figure 5:33 The neural specific fragment of the Tet3 upstream putative promoter is highly conserved to the human sequence. Alignment of the murine Tet3 upstream neural-specific putative promoter fragment (co-ordinates: +2021 to +2184) with the human sequence are shown. The mouse sequence is highlighted in bold. * indicates sequence homology. Grey shading highlights a putative binding site for the Snal1 TF, see section 5.3.19. Numbers correspond to the number of bps from the most 5' cap site (+1), see Figure 5.10.

5.3.19 Snal1 was identified as a potential transcription factor responsible for the neural-specific effects of the Tet3 downstream and Tet3 upstream putative promoter fragments

To begin to elucidate potential TFs that may be responsible for mediating the neural-specific cis-acting transcriptional regulation, identified in both Tet3 upstream and downstream putative promoter fragments, potential TF binding sites were assessed in these regulatory sequences using the bioinformatics tool Consite (http://consite.genereg.net). The potential neural regulatory sequence at the co-ordinates -493 to -110 and +450 to +1441 in the Tet3 downstream putative promoter (see Figure 5:22 and Figure 5:32) and +2021 to +2184 in the putative Tet3 upstream promoter region (see Figure 5:30 and Figure 5:33) was found, among a selection of TFs, to have 7, 13 and 1 Snail family transcriptional repressor 1 (Snal1) binding sites respectively. Snal1 is key factor involved in embryogenesis and is strongly implicated in neural crest formation[519,520,521]. To begin to investigate the potential regulation of Tet3 by Snal1, the mRNA expression of Snal1 was assessed over a time course mESC differentiation (Figure 5:34). The expression profile of Snal1 was similar to Tet3 (Figure 4:5), whereby an induction of mRNA was evident after day 3 of differentiation. This suggests the possibility that Snal1 could regulate Tet3 mRNA expression, although further experimentation is required to determine this.

It can also be noted TF binding site analysis revealed the ectodermal factor Sox5[522] has 8 putative sites within the neural-regulatory sequence in the Tet3 downstream promoter fragment, at the co-ordinates +450 to +1441 (Figure 5:32B).

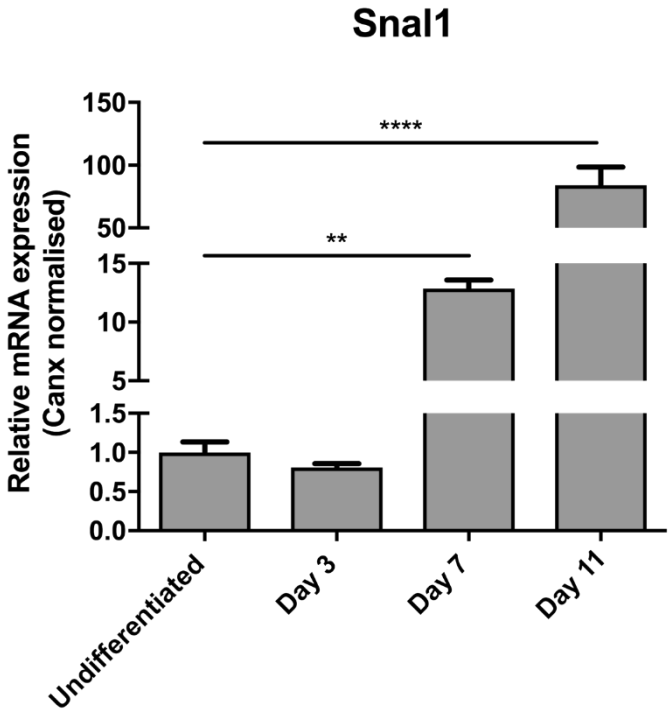


Figure 5:34 Snal1 expression is increased over a time course of mESC differentiation. The expression of Snal1 was assessed by QPCR over a time course of mESC differentiation. Data are expressed as the mean ± SEM relative to undifferentiated cells, and analysed by a one-way ANOVA with Dunnett's post-hoc analysis. ** $p<0.01$, **** $p<0.0001$

5.4 *Discussion*

5.4.1 5-hmC is specifically expressed in mouse brain compared to heart

Differential levels of 5-hmC in brain tissue, compared to heart, at 2 developmental stages (E 18.5 embryos and 1-week old pups) was confirmed by 2 analytical methods, dot-blot and mass-spectrometry (Figure 5:2, Figure 5:3 and Figure 5:4). Consistent with the literature[260,259,523], the data shown here demonstrates higher levels of 5-hmC in brain tissue, consistent with this base having a functioning role(s) in the mammalian brain[524]. The precise function of 5-hmC in the brain is unknown, however observational studies have shown 5-hmC is localised to promoter regions and more abundantly within gene-bodies, correlating with high transcriptional activity[525,182,506]. How deposition of 5-hmC increases transcription is unknown, however it can be speculated that the DNA destabilisation effect of this polar 5-hmC mark facilitates TF binding to promote gene expression[526,527].

5.4.2 Tet3 is involved in neuronal cell function and differentiation

Tet3 mRNA expression was found to potentially account for the increased 5-hmC in brain, compared to heart (Figure 5:5 and Figure 5:6), coinciding with previous studies that have reported Tet3 as the predominant paralogue in neuronal tissue[422,506,518]. However, Tet3 expression only trended to increase in embryonic brain above heart, suggesting the 2.5–fold induction in the 5-hmC:5-mC ratio may be attributed to a higher level of Tet protein expression and/or activity. However, in 1-week old pups Tet3 mRNA was significantly elevated in brain, compared to heart tissue, indicating a potential cell-specific role of Tet3.

Literature evidence suggests Tet3 is involved neurogenesis[321,513]. To further investigate this Tet3 expression was determined in P19 cells directed to differentiate into neural or cardiac-like cells in the presence of RA or DMSO respectively[528,529]. Tet3 mRNA showed a dynamic profile in RA-treated cells over the tested 10-day time course (Figure 5:9), expression levels showed a time-dependent increase till day 6, which then persisted to decrease to levels equivalent to DMSO-treated cells at day 10. This may further provide supporting evidence for cell-specific function of Tet3. Consistent with these findings, it has previously been shown in mESCs that RA, compared to DMSO treatment, induced Tet3 expression (while decreasing Tet1 and Tet2 mRNA levels) to promote 5-hmC formation[530]. In addition, an independent study into neurogenesis revealed Tet3 expression was found to be low in undifferentiated ESCs and mature neurons, but high

in neuronal progenitors and immature neurons, thereby supporting a role of Tet3 in neurogenic events[531]. Intriguingly, from the data shown here (Figure 5:9) Tet2 expression closely reflected the Tet3 profile, but in fact achieved a higher peak expression at day 6 (approximately 8-fold for Tet2, compared to 4.5-fold for Tet3) in RA-treated cells. Tet2 has also been linked to neural processes, specifically interacting with the histone methyltransferase AF9 to facilitate neurodevelopmental gene activation[532] and neuronal-cell survival[34]. However, Tet2 was also induced in DMSO-treated P19 cells, above undifferentiated levels after day 6 of differentiation. Hence this infers that Tet2 may have a developmental role in more than one cell type.

5.4.3 Tet3 has 2 protein isoforms that potentially may have differential roles in neuronal cell differentiation and function

To begin to elucidate the regulation of Tet3, which may underlie lineage-specific functions, 5'RACE was performed on day 7-differentiated mESCs samples to identify potential transcriptional start sites. Consistent with previous studies[515,175], multiple cap sites were located upstream of 2 ATG start codons, at 2 distinct regions, encoding a longer CXXC-containing variant (Tet3 uptream) and a shorter alternative that lacks this domain (Tet3 downstream) (see Figure 5:10).

The exact role of these variants in neuronal cell differentiation and function are controversial[515,175,516] (see section 5.1.2), and thus attempts were made to decipher the potential differential roles of these variants. The Tet3 downstream variant specifically demonstrated increased expression in brain, compared to heart, in 1-week old pups (Figure 5:13). Despite not being fully mature at this age, mice have undergone a level of neurogenesis[533]. Further, the Tet3 downstream transcript showed sustained expression, compared to the Tet3 upstream, in neural over cardiac-like directed P19 cell differentiation (Figure 5:14). However, it should be noted the maturity of neurons derived from directed P19 cell differentiation is unknown. In comparison, the Tet3 upstream transcript showed a transient induction in mRNA expression in neural-directed P19 cells, relative to undifferentiated levels (Figure 5:14). This was found to be consistent with the immediate induction of this transcript at the onset of unbiased mESC differentiation (Figure 5:15 and Figure 4:31), which correlated with subsequent neuronal maker induction (Figure 3:9). In addition, the loss of neuronal progenitor expression in 1% O_2-exposed differentiating mESCs correlated to a greater loss in expression of the upstream, compared to the downstream transcript. (Figure 5:16). Therefore, these distinct patterns of mRNA expression suggest the 2 protein isoforms may have independent roles, and hence it is tempting to speculate from the albeit limited evidence

shown here that the downstream transcript may be linked to a neuronal maintenance role, whereas the upstream to a more predominant role in early neurogenesis.

5.4.4 Neural-specific regulatory fragments were identified in the both the Tet3 downstream and Tet3 upstream putative promoters

The putative promoter regions of the 2 distinct protein coding transcripts (Tet3 upstream and downstream) were cloned into luciferase reporter constructs as tools to investigate the transcriptional regulation that may underlie the potential independent functions of these protein isoforms.

Initial assessment of reporter activity in early differentiating P19 cells and mESCs demonstrated consistently higher activity in the cloned putative promoter fragment of the downstream, compared to the upstream region (Figure 5:18 and Figure 5:19), which may thus be reflective of the mRNA expression profiles of these transcripts at the onset of mESC differentiation (Figure 5:15). In addition, the cloned downstream fragment displayed higher promoter activity, relative to the upstream promoter in SH-SY5Y cells (Figure 5:20). Although it should be considered that SH-SY5Y cells are not reminiscent of fully mature neurons[534,535], 5' and 3' deletions of this downstream promoter fragment was performed to identify regions of sequence conferring potential neural-specificity (Figure 5:22 and Figure 5:25).

Analysis of potential TF binding sites (using the Consite bioinformatics tool) within the identified neural-regulatory sequences of the Tet3 downstream promoter fragments (Figure 5:32) identified 7 and 13 putative Snal1 binding sites within the co-ordinates -493 to -110 and +450 to +1411, relative to the number bps from the most 5' cap site. Additional analysis revealed 8 Sox5 binding sites within the co-ordinates +450 to +1411 (Figure 5:32B), which were more highly conserved with the human sequence than the Snal1 sites. Sox5 has previously been reported to be intrinsic to ectodermal formation[522], involved in oligodendrocyte[536] and corticofugal neuron[537] development. Therefore, this provides a basis for future investigation into a possible role of Sox5 mediated regulation of Tet3 mRNA.

Owing to a large defined potential promoter region encompassing the Tet3 upstream variant, on the Ensembl database, it was speculated that an additional longer regulatory sequence than that originally cloned for the Tet3 upstream putative promoter may exist. A 3' extension of the cloned Tet3 upstream promoter, named extension A, (Figure 5:27) was found to contain a positively regulatory sequence (Figure 5:28), which demonstrated

neural-specificity (Figure 5:29). Consecutive deletions from extension A identified a potentially neural-regulatory sequence of only 163 bps (Figure 5:31). Analysis of putative TF binding sites at the co-ordinates +2021 to +2184, relative to the number bps from the most 5' cap site, identified a Snal1 putative binding site (Figure 5:33) that is conserved with the human sequence. The Snail family has an intrinsic role in morphogenesis, but is specifically linked to promoting cell movements such as in the process of EMT and neural crest formation[520,538,539]. Snail family members are considered as transcriptional repressors, functioning to block E-cadherin expression (enabling EMT) and maintain germ layer boundaries *via* repression of ectodermal genes within the mesodermal domain[540,541]. However, there is also evidence for Snail to potentiate transcription[542,543,544], which supports *in vivo* data demonstrating a role of Snail as a key regulator of neural precursor cells in embryonic and adult mice[521]. Therefore, considering Snal1 and Tet3 have similar gene expression profiles during mESC differentiation (Figure 5:34 and Figure 5:15), it can be proposed that Snal1 is a potential candidate for regulation of Tet3 expression. Supporting this hypothesis, depletion of Tet3 in xenopus resided in a loss of Snal1 expression[176], suggesting these factors are intrinsically linked, however further investigation is required to determine any functional regulation between them.

Overall, the luciferase reporter data shown here goes some way towards identifying novel neural-specific regulatory sequences within the Tet3 promoter region, which can be used for future investigations to identify TFs that may regulate Tet3 in neurogenic events.

5.4.5 Luciferase reporters are unlikely to be under epigenetic control and thus data should be interpreted with caution

Perhaps surprisingly, the activity of the Tet3 downstream putative promoter construct was elevated in undifferentiated P19 cells (Figure 5:21), whereas the mRNA expression of the Tet3 downstream transcript, albeit in differentiated mESCs, was decreased (Figure 5:16) in 1% O_2. Despite an inability to express luciferase reporters transiently for a long enough period to study the effects of [O_2] in differentiated cells, evidence from Chapter 4: Results 2 demonstrated that the Tet3 downstream promoter region showed O_2-dependent enrichment of 5-hmC in day 3-differentiated mESCs (Figure 4:27). This therefore demonstrated the Tet3 downstream promoter is under epigenetic control. However, transient transfections of DNA plasmids are devoid of *de novo* methylation (and thus hydroxymethylation) since they are not replicated. Evidence in the literature has only reported such an event in the instance of replication of an artificial chromosome-

like vector over a larger period of time[545]. Therefore, data achieved from luciferase reporters should be interpreted with caution, as the effect of epigenetic regulation are unlikely to be recapitulated in the transiently-transfected plasmids.

5.4.6 Conclusion

In conclusion, the data shown here confirms 5-hmC levels and Tet3 mRNA expression are elevated in brain, compared to heart tissue. To understand how Tet3 governs potential neural-specificity during development, 2 known Tet3 transcript variants, which encode distinct protein isoforms, were firstly confirmed in differentiating mESCs. Assessment of mRNA expression and putative promoter activity of these transcript variants suggest they are under differential regulation, and therefore may have not only cell-type specific functions, but also distinct roles in neurogenesis. Finally, the identified highly conserved neural-specific regions in both the Tet3 upstream and downstream promoter fragments provide a future basis to confirm whether the potential TFs Snal1 and Sox5 are responsible for Tet3 regulation.

Chapter 6: General Discussion

6.1 *Summary and proposed mechanism of O_2-dependent regulation of cell fate decisions*

The data presented here demonstrate for the first time that the Tet enzymes are differentially regulated by [O_2]. Specifically, Tet1 was identified as the paralogue most likely to be differentially inhibited at O_2 tensions considered physiologically relevant within the developing mammalian embryo, thereby presenting as an O_2 sensor. It is proposed that Tet1 regulates Tet3 expression in a spatial manner (dependent upon an adequate O_2 supply) to in part control cell fate decisions, depicted in the schematic shown in Figure 6:1. In undifferentiated mESCs, Tet1 is localised to a CG rich region of the Tet3 promoter, which is also marked by active (H3K4me3) and repressive (H3K27me3) histone marks, characteristic of a bivalent promoter. Global levels of 5-hmC (and 5-mC) are low in undifferentiated mESCs, and the Tet3 promoter is devoid of these marks when cells are in this pluripotent state. Upon differentiation, it is proposed that Tet1 activity is increased in a graded manner, dependent upon O_2 availability, to induce hydroxylation at a CpG rich region of the Tet3 downstream promoter, which encodes a truncated protein isoform, lacking a CXXC DNA binding domain. Concomitant with the loss of the H3K27me3 and maintenance of H3K4me3, Tet3 becomes actively transcribed to potentially promote neuronal cell specification. Consistent with previous reports[515,175], transcripts encoding 2 distinct Tet3 protein isoforms with and without a CXXC binding domain were identified in differentiated mESCs and tissue samples *in vivo*. These transcripts were shown to be differentially regulated by distinct promoter regions that may underlie independent functions in developmental events, but requires further elucidation.

Overall, these findings provide a novel mechanistic insight into how the regulation of epigenetic modifications *via* microenvironments of O_2 within the developing embryo can function to control cell fate decisions. The context and considerations of these findings in relation to the literature are discussed below.

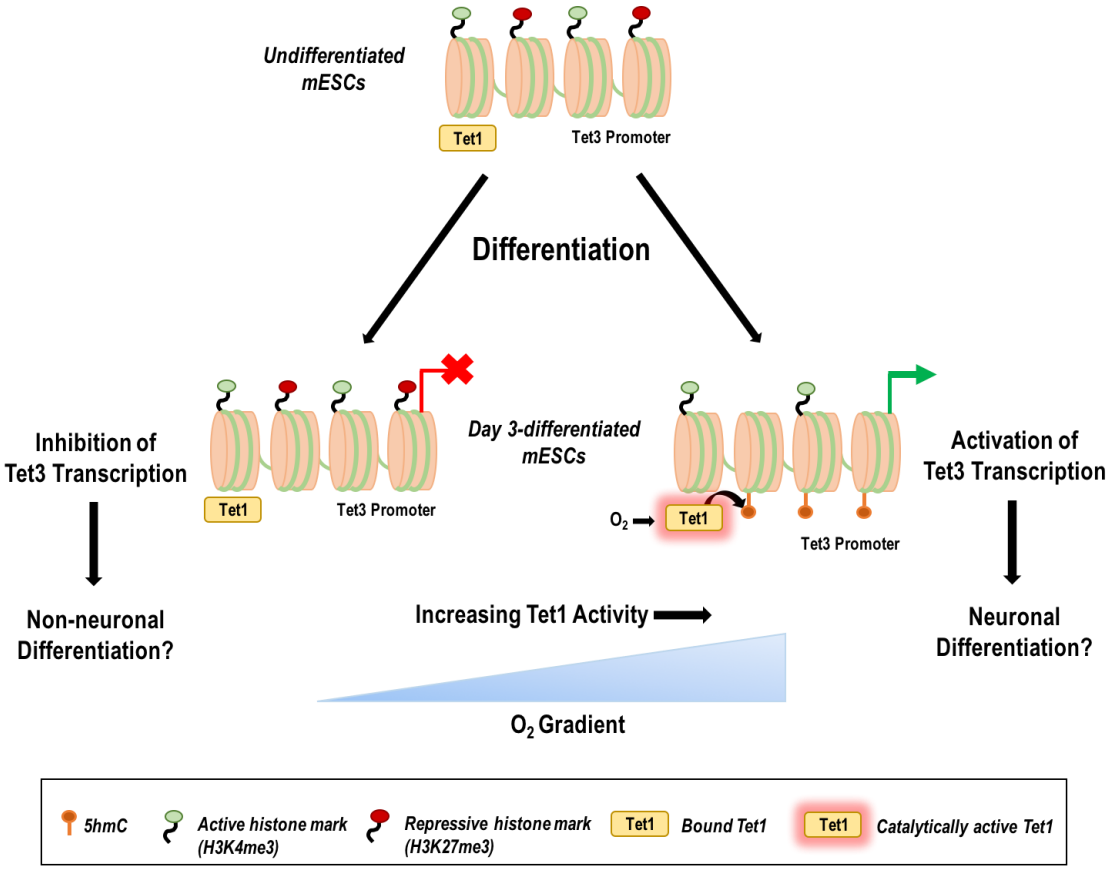


Figure 6:1 Schematic representation of the proposed O_2-dependent regulation of Tet3 mRNA expression during cellular differentiation. In undifferentiated mESCs Tet1 is bound to a CG rich region of the Tet3 promoter, marked with the active (H3K4me3) and repressive (H3K27me3) histone marks. Upon differentiation, Tet1 activity increases in an O_2-dependent manner to generate 5-hmC within the Tet3 promoter region. Concomitantly, H3K27me3 is lost and H3K4me3 is maintained, which activates Tet3 transcription to potentially promote neural differentiation pathways. Image adapted from Burr, S *et al*. Nucleic Acids Res. 2017; ahead of print[481].

6.2 *O_2 signalling in the embryo is likely to encompass a multitude of pathways to contribute to developmental gene regulation*

The data shown here demonstrated that the morphogenic properties of O_2 in early cellular differentiation may be attributed at least in part to the differential sensitivities of Tet enzymes to O_2 tensions, thereby suggesting a possible way to drive asymmetry within the early embryo. However, the effect of genetic ablation of Tet enzymes on cellular specification produced highly variable results on lineage marker expression, and thus failed (with the exception of pluripotency markers) to recapitulate the skewing of

mESC differentiation upon exposure to 1% O_2. This raises 2 possibilities; Tet enzymes may show functional compensation and/or redundancy, meaning that the precise function of each paralogue is difficult to delineate (discussed in section 6.5) and/or the involvement of other O_2-dependent signalling pathways. An additional epigenetic regulatory mechanism, which controls the methylation status of histones, involve the 2-OGDO domain containing JmjC-KDMs proteins. The ability of this 7-member demethylase family to function as a potential O_2 sensor remain to be elucidated, however initial K_m studies performed on the KDM4 subtype suggest these enzymes are sensitive to physiological O_2 tensions in the developing mammalian embryo[273,274]. From studies performed here, the global protein expression of the histone marks H3K27me2/3 and H3K4me3 were elevated upon exposure of 1% O_2 in mESCs, which may be indicative of inhibition of KDM6 and KDM5 respectively. Thus, owing to the presence of these 2 marks at bivalent promoters, it may be speculated that these JmjC-KDMs may function similarly to Tet proteins in a spatial manner, based upon O_2 availability, to control developmental gene expression and thus asymmetry in the embryo.

It should also be considered that O_2, in addition to regulating the activity of these enzymes, may also affect their transcriptional regulation. For example, data attained from these studies demonstrated that exposure of differentiating mESCs to 1% O_2 acted to elevate Tet3 mRNA expression at day 3, but by days 7 and 11 mRNA levels were decreased, compared to atmospheric O_2. A recent study using *in vitro*-produced bovine embryos showed elevated Tet1 and Tet3 mRNA expression in 5%, compared to atmospheric O_2[546]. Previously, it has been shown that culture of neuroblastoma cells in 1% O_2 induced HIF-1-dependent Tet1 mRNA expression, which was subsequently demonstrated to deposit 5-hmC at hypoxia-responsive gene promoters to induce transcriptional activation[495]. Similarly, ablation of HIF-1α in HepG2 cells attenuated the 1% O_2-induced expression of Tet1, Tet2 and Tet3[547]. The potential for HIF-dependent regulation of Tet3 was highlighted from the luciferase reporter data in P19 cells from this study, whereby activity of the Tet3 downstream promoter, harbouring 8 core HIF DNA binding motifs, was elevated in 1%, compared to atmospheric, O_2. However, it should be noted, as previously referred to in section 4.4.2, the effects of [O_2] upon Tet activity and transcriptional regulation is cell-type dependent[309]. Similarly, KDMs have been reported to be under the transcriptional control of HIFs[269], and intriguingly demonstrate differential specificity, whereby HIF-1α induced KDM3A and KDM4B[548] and HIF-2α induced KDM6A mRNA expression[549].

Overall, the ability of graded levels of O_2 to influence the dynamics of DNA and histone methylation may provide an insight into mechanisms regulating transcriptional programs

in the developing embryo. The fact that HIF signalling appears to be a regulator of these epigenetic modifications at the transcriptional level may further indicate its intrinsic role as a master regulator of O_2 homeostasis[550]. This may function as either a compensatory mechanism for the inhibition of 2-OGDO activity and/or provide an additional mechanism to ensure precise control of gene regulation at the epigenetic level. However, it is also plausible to imagine that additional O_2-dependent mechanisms regulating cell fate decisions may exist in what is a highly co-ordinated, complex and relatively unexplored field.

6.3 *Tet1 was identified as a potential O_2 sensor in the developing mammalian embryo but also may have non-catalytic functions*

From ectopic expression of Tet1, Tet2 and Tet3 in HEK cells, exposed to a graded set of low O_2 levels, it was concluded that Tet1 activity specifically was inhibited by physiologically relevant O_2 tensions. Currently, the concept of O_2-dependent regulation of Tet activity in differentiating mESCs is unexplored in the literature. The only study currently that has hinted at the ability of Tet1 to mediate a hypoxic response in mESCs arose from the exposure of undifferentiated cells to 0.5% O_2, which was shown to have a lesser inhibitory effect on 5-hmC levels in Tet1-knockout, compared to wild-type cells[309]. The significance of Tet1 in O_2-dependent signalling was demonstrated recently from a series of *in vivo* studies. Tet1-null zebrafish showed lower survival rates upon 5% O_2 exposure compared to Tet2 and Tet3 silenced fish[551]. Similarly, it was shown that Tet1 ablation in mice showed lower survival rates upon exposure to 10-8% O_2, compared to wildtype litter mates[551]. Thus Tet1 ablation appeared to impair hypoxia tolerance, perhaps consistent with Tet1 activity being the most sensitive to changes in [O_2] as demonstrated here. Interestingly, this study reported Tet1 could enhance the transcriptional activation of HIF-1α through a stabilising interaction with the C-terminus, and of HIF-2α, *via* competing with PHD2 for HIF-2α binding to reduce hydroxylation[551]. This provides a unique insight into Tet1 function, which is independent of its enzymatic activity.

An additional non-catalytic role of Tet1 has been highlighted to be responsible for a non-redundant role at the post-implantation stage of embryogenesis. Tet1 was found to repress Epi target genes independent of methylation changes *via* Jmjd8[497]. Jmjd8, similar to the founding member of the Jmj family Jarid2, lacks essential residues necessary for catalytic activity, but functions as a transcriptional repressor[497,552]. While the catalytic domain of Tet1 was found to be dispensable for development, in agreement

with previous studies[251,553], deletions of the N-terminal fragment downregulated the Jmjd8 gene and was associated with gastrulation defects resulting lethality[497].

The data from this study found Tet1 to be bound to a CG rich region of the Tet3 promoter, encoding the protein isoform lacking the CXXC domain in undifferentiated and day 3-differentiated mESCs. The function of Tet1 at this site may be twofold. Tet1 has previously been shown to be localised to CpG regions at promoters of actively transcribed genes and Polycomb-repressed genes. In ESCs, Tet1 can facilitate gene repression through a complex with Prc2 and thus association with the H3K27me3 mark[194,456]. Furthermore, Tet1 binding at gene promoters was also found to overlap with an additional histone mark, H3K4me3, and thus Tet1 is therefore intrinsic to bivalency (a chromatin state characterised by the presence of repressive (H3K27me3) and activating (H3K4me3) marks)[194]. An insight into the association of Tet1 with these marks was highlighted from assessment of global protein levels, which confirmed that specific silencing of Tet1 induced H3K4me3 and for reasons currently unknown the di-methyl form H3K27. For the first time, the presence of these histone marks were identified at the same Tet3 promoter region as Tet1 in undifferentiated mESCs. Therefore, it was concluded that Tet3 is characteristic of a bivalent promoter, and hence can be considered a developmental gene poised for activation.

In agreement with the literature[554], this bivalent domain was also found to be hypomethylated, which may be attributed to the low catalytic activity of Tet1, or as shown here, more likely, due to the low levels of 5-mC in undifferentiated ESCs, compared to differentiating cells. When induced to differentiate, by day 3, this Tet3 promoter region showed a loss of the H3K27me3 mark (H3K4me3 remained) and a gain of hydroxylation, correlating with the induction of Tet3 mRNA expression. Although this was not shown directly, the data achieved from the studies performed here strongly supports that the induction of O_2-dependent 5-hmC at this site was primarily mediated by Tet1. Silencing of Tet1 failed to reduce 5-hmC at the Tet3 promoter region, potentially attributed to Tet compensatory effects (discussed in section 6.5). Consistent with this hypothesis, in mESCs it has been reported 5-hmC is preferentially enriched at Tet1-bound gene promoters and intragenic regions[390]. But, upon Tet1 depletion 5-hmC is mainly lost from intragenic regions, suggesting potential functional redundancy between Tet1 and Tet2 at Tet1 bound gene promoters[390,506]. Therefore, the positioning of Tet1 on the Tet3 promoter region facilitates the target specificity of 5-hmC, and therefore suggests that the pre-requisite formation of 5-mC mark is closely coupled to hydroxylation. In light of this, it has been reported that hypermethylation of bivalent promoters is associated with transcriptional activation[256], consistent with notion that 5-mC impairs recruitment of

Prc2[454]. Thus, with respect to Tet3 regulation specifically, it can be hypothesised that initial methylation events may be required to remove the H3K27me3 mark before subsequent oxidation by Tet1 to generate 5-hmC. However, the mechanism(s) underlying the induction of Tet1 catalytic activity remains to be elucidated, but was confirmed not to be a result of AA-dependent modulation of Tet activity.

Collectively, these investigations support the dual functionality of Tet1 at the Tet3 promoter; a non-catalytic role for the formation of a bivalent domain and the catalytic enrichment of 5-hmC.

6.4 *Tet1 and Tet3 are implicated in neurogenesis*

The investigations performed here reveal a novel mechanism highlighting the potential interdependence between Tet1 and Tet3, which may have implications in cell fate decisions. Tet3 is the most abundant Tet paralogue in neurons[518,506] , and has been documented to be fundamental in neural development[321,322] as well as the transdifferentiation of fibroblasts to functional neurons[555]. The regulation of neural-related genes, such as Pax6, Ngn2 and Tubb2b is thought to occur by the co-ordination of the CXXC binding domain and the intrinsic hydroxylase activity of Tet3[176]. Consistent with this notion, 5-hmC levels has shown to be increased during neurogenesis, and associated with higher transcript levels of genes critical for neuronal differentiation, migration and axonal guidance[506]. Interestingly, in contrast to the negative cross-talk reported between 5-mC and H3K27me3 in neural stem cells (loss of 5-mC resulted in accumulation of H3K27me3)[454], a negative association of 5-hmC with H3K27me3 was reported in neuronal differentiation[506]. This observation is consistent with the data reported from day 3-differentiated mESCs at the Tet3 promoter, in which the gain of hydroxylation is associated with the loss of H3K27me3.

Hydroxylation at the Tet3 promoter was lost in 1% O_2, which correlated with transcriptional inhibition in day 7 and 11-differentiated mESCs. Associated with these findings was the loss of neuroectodermal marker expression (Figure 3:11 and Figure 3:15). Therefore, it can be speculated that O_2-dependent regulation of Tet3 transcription can function to regulate, in part, formation of the ectodermal lineage. The data reported here suggests that O_2-dependent activity of Tet1 may regulate isoform-specific Tet3 expression through deposition of the 5-hmC mark. An insight into the possible functional co-operation between Tet1 and Tet3 was suggested from expression of both paralogues appearing at the 2-cell stage of embryogenesis (Tet3 expression is however rapidly lost after the first cleavage events, re-appearing later during cellular differentiation)[556]. The

requirement of both paralogues in early embryonic development were highlighted from double deficient mice, whereby no viable embryos were detected past E 10.5, which was attributed to poor forebrain formation[556]. In further support of this, Tet1 has also been implicated in brain function[557]. Specifically, Tet1-ablated mice induced hypermethylation, corresponding to decreased expression of neuronal-activity regulated genes, resulting in an impairment of extinction learning[558]. These deleterious effects of Tet1 silencing were confirmed independently and additionally were shown to result in decreased adult neural progenitor cell proliferation, resulting in impaired neurogenesis[553].

The requirement of Tet3 in neuronal differentiation was recently demonstrated through loss and ectopic expression of Tet3 in a mESC differentiation model[322]. This effect was not recapitulated in the Tet3 silencing studies performed here (Figure 3:30), likely due to differences in the method of cellular differentiation. Li, X *et al* (2016)[322] implemented neural-restricted differentiation conditions, compared to the unbiased differentiation approach used here. This was reflective of a more pronounced induction of Tet3 mRNA expression, recorded to peak at approximately 40-fold[322], compared to 7-fold, under neural and unbiased differentiation conditions respectively. In addition, discrepancies may also be attributed to Tet functional redundancy (see section 6.5) and differences in the Tet3 deletion strategy. Li, X *et al* (2016)[322] utilised complete genetic knockout of Tet3, whereas an shRNA-mediated gene silencing approach was used here (albeit an efficient level of silencing, ~90%, was achieved). Aside from the impaired neural specification, seen with Tet3 ablation, a skewing of differentiation towards the cardiac mesoderm was also reported[322]. Thus inhibition of Tet3-mediated signalling may also be required to mediate mesodermal differentiation, suggestive of Tet3 acting as a switch to control lineage specification.

The potential cell specificity of Tet3 was considered in the context of its 2-distinct protein-coding isoforms, which were confirmed to be present in differentiated mESCs and tissues *in vivo*. Evidence from these investigations suggests Tet1 may regulate isoform specific Tet3 through deposition of the 5-hmC mark at the Tet3 downstream promoter region, which drives expression of the smaller truncated protein, lacking the CXXC domain. However, it should be noted Tet1 silencing failed to decrease the mRNA expression of this Tet3 downstream transcript (and/or the Tet3 upstream variant). This may be attributed to the inability of Tet1 silenced mESCs to reduce hydroxylation at this Tet3 promoter, potentially caused by either Tet functional compensation (see section 6.5) or a lack of sufficient Tet1 ablation, enabling some functionality to remain. However, the exact function of 5-hmC at this region is unknown and will form the basis for future investigations. In light of the controversy in the literature surrounding the neuronal

involvement of either Tet3 isoform[515,175,516], evidence from Chapter 5: Results 3 demonstrate these protein coding transcript variants confer differential regulation and therefore suggests they may have distinct functions in neurogenesis.

6.5 *Functional compensation and redundancy between Tet proteins mask their independent developmental roles*

The precise function of each Tet paralogue in cellular differentiation may be difficult to elucidate due to compensatory and functionally redundant effects between the Tet enzymes. This was shown to be apparent not only in the undifferentiated stable Tet silenced cell lines generated for use in this study, but also widely reported in the literature[248,251,258,322]. For example, recently it was shown that depletion of Tet3 in embryos, compared to Tet1/2/3-deficent embryos, demonstrated no overt neuronal defects, which was suggested to be caused by functional redundancy from Tet1 and Tet2[322]. In another instance, due to compensatory mechanisms, developmental defects in gastrulation required deletion of all 3 Tet paralogues[258]. A final study important in portraying this phenomenon occurred when depletion of Tet1 in mESCs (the most highly expression paralogue in these cells) led to only a partial reduction in 5-hmC, leading to suggestions other Tet family members could compensate for this loss[251]. However, importantly no change in mRNA levels of either Tet2 and Tet3 was confirmed[251], leading to speculation that perhaps the catalytic activity of these paralogues can be induced.

6.6 *Considerations and Future directions*

ESCs are a widely-used tool for studying and modelling cell specification during cellular differentiation events. ESCs, derived from the ICM, can be maintained in a ‘ground state’ of pluripotency (devoid of methylation) in the presence of 2i and LIF culture conditions[349]. Differentiation of these cells form primed Epi-like cells, mimicking the E 5.5 – E 6.0 pre-gastrulation Epi (but do not form the TE in culture), and thus are suitable for investigating early embryogenesis[559]. The formation of EBs, as performed in these studies, provide a 3-dimensional structure that enhances cell-cell interactions, which may be of importance to specific developmental programs[560]. However, EBs typically form different sizes cell aggregates and an un-patterned heterogeneous population of cell derivatives. Therefore, the spatial organisation of the EB is unlikely to reflect that of the developing early embryo, and thus future investigations could encompass advances in differentiation strategies aimed at recapitulating developmental patterning[561,562,563]. In addition, the use of iPSCs, specifically derived from human, may provide a more

translational approach for the investigation of Tet enzyme function, with the aim of understanding their role(s) in cell specification events to facilitate generation of patient specific cells for transplantation. Research has already implicated Tet proteins and 5-hmC in the epigenetic reprogramming required for generation of iPSCs[564], and in fact Tet1 has been shown to be able to replace the Oct4 in somatic cell reprograming[565]. However, upon assessment of the current literature there is no evidence that Tet proteins have been investigated in iPSC directed cellular differentiation.

In light of the current work performed here, demonstrating the potential involvement of [O_2] and Tet proteins to dictate asymmetry and cell fate decisions in an unbiased model of cellular differentiation, it is of potential importance to assess the role of 5-hmC/Tet proteins in specific cellular differentiation programmes. Considering it has been previously reported that Tet3 may function as regulator of both mesoderm and ectodermal differentiation[322], the utilisation of well documented methods to direct ESCs to cardiac[566,567,568] and neural lineage[569,570,571] may prove useful to gain a mechanistic insight into the exact cell specificity of the Tets. Directed differentiation strategies, in combination with the use of CRISPR/cas9, for the generation of a complete Tet knockout pluripotent cell line (which could be performed on individual a Tet, or in combination to produce double or even triple Tet ablated cells), would provide a valuable tool for elucidating the exact functional role of each Tet in a defined cellular specification pathway. Furthermore, a CRISPR-based approach could be implemented to target Tet enzymes to specific gene loci to induce DNA demethylation[572], hence this approach could be implemented during ESC differentiation, to control gene expression, and thus facilitate derivation of cells for transplantation.

A directed ESC differentiation approach could be additionally implemented to delineate between the 2 protein coding isoforms of Tet3 in neuronal specification. One proposition may be to generate stable ESC lines expressing the luciferase reporters driven by the Tet3 upstream and downstream promoter fragments. This would enable firstly the possibility of detecting promoter activity at the time point Tet3 mRNA expression is induced during differentiation and provide information into which promoter is more active during neurogenesis. The effect of each isoform on neuronal differentiation can then be addressed using a genetic ablation strategy and assessing mRNA expression of key early (Dcx[354], Sox2[573]) and late (Map2[347], NeuN[574]) neuronal markers. Additional insights into the functional roles of these isoforms could be conducted *in vivo*, whereby staining for each isoform in the mouse brain at a variety of developmental stages (embryonic to adult) may confirm differential roles in neurogenesis and neuron function. A final follow up to this work is to identify TFs that regulate expression of these Tet3 transcript variants.

Recently it was identified NeuroD1 could induce Tet3 pan expression[575], and in combination with the bioinformatics analysis, which identified potential binding sites for Snal1 and Sox5 within the Tet3 promoter region, provide candidate TFs to explore. To confirm this potential regulation of Tet3, overexpression and genetic ablation of these factors in neural and cardiac cell types can be performed to assesses an ability to potentiate promoter activity and Tet3 expression. Further verification can be made by mutagenesis studies on the proposed TF binding sites on the Tet3 promoter, followed by subsequent assessment of promoter activity in neural and cardiac cells.

Building from these studies, future experimentations can be targeted towards an *in vivo* approach. To verify Tet1-dependent regulation of Tet3, which may be of importance in neurogenesis, Tet1-knockout mice, or potentially Tet1-conditional knockout mice, under control of a neuronal progenitor specific Cre-recombinase such as Nestin[576] or Emx1[577], can be used to assess Tet3 mRNA expression as well as Tet3 staining in the brain. This system could also be implemented within an O_2 controlled environment, whereby mice heterozygous for a deletion in Tet1 can be mated and subsequent pregnant mice maintained in mild hypoxia conditions (8% O_2) (lowest level of tolerance was determined as 5.5% O_2), similarly to studies performed by Sparrow, D *et al* (2012)[578]. Pregnant mice would be sacrificed at varying stages of early embryonic development and the effects upon neuronal differentiation assessed. This study would provide information on the impact of neuronal development when exposed to low O_2 tensions in the absence of the *in vitro* determined Tet1 O_2 sensor. Furthermore, as similarly performed from mESC differentiation studies, it can be verified whether O_2-dependent Tet1-mediated hydroxylation on the Tet3 promoter occurs in the mouse brain. The study of Tets and O_2 in embryogenesis may be further examined through whole embryo culture. Utilisation of previously documented methods[579,580] would enable the exposure of Tet deficient embryos to graded levels of low O_2 tensions, thereby assessing the O_2 sensing capabilities of Tets in a more physiological model compared to *in vitro* cellular differentiation. This would further provide insights into the effects of Tet ablation and [O_2] on developmental patterning and asymmetry. In addition, genomic approaches can be implemented here to identify the effects of [O_2] on DNA demethylation within the developmental genome, and thus provide information on gene regulation, which will increase the current understanding of embryogenesis.

In consideration of the studies shown here, some final points should also be mentioned that may provide other avenues for future exploration. Firstly, the work performed here focused only on the 5-hmC mark, however literature evidence suggests that despite being a rare base in the mammalian genome, the 5-fC modification is also stable and

has different dynamics to 5-hmC, thereby suggesting this mark may have an independent functional role[581]. Genome-wide mapping data revealed a greater state of 5-fC enrichment at CpG islands of promoters than 5-hmC, and interestingly found a correlation with the H3K4me3, suggestive of a role in transcriptional activation[582]. Therefore, owing to differential substrate specificities of the Tet enzymes[198,206], it could be hypothesised that changes in [O_2] may have less impact upon 5-fC levels and thus may have consequential significance upon gene regulation. The complexity of gene regulation is further evident from the fact that RNA molecules themselves can become methylated to form 5-methylcytidine (5-mrC) and N^6-methyladenonsine (m^6A), which is reported to have influential role on RNA stability, localisation and translation[583,584]. Intriguingly, Tet enzymes can oxidise 5-mrC to form 5-hydroxymethylcytidine, whereas the m^6A mark can be oxidised to N^6-hydroxymethyladenonsine (hm^6A) by AlkB family of dioxygenases (including fat mass obesity associated isoform)[585,586,587]. Similarly to the Tets, the activity of these enzymes are also dependent upon [O_2][586], and is of potential interest for prospective investigations to assess the sensitivity of this class of enzyme to graded levels of low O_2 tensions.

References

1. Mason C, Dunnill P. A brief definition of regenerative medicine. *Regen Med*. 2008;3(1):1-5. doi:10.2217/17460751.3.1.1.
2. Caplan AI, Mason C, Reeve B. The 3Rs of Cell Therapy. *Stem Cells Transl Med*. 2017;6(1):17-21. doi:10.5966/sctm.2016-0180.
3. Ellis H. James Blundell, pioneer of blood transfusion. *Br J Hosp Med*. 2007;68(8):447-447. doi:10.12968/hmed.2007.68.8.24500.
4. Laflamme MA, Murry CE. Heart regeneration. *Nature*. 2011;473(7347):326-335. doi:10.1038/nature10147.
5. Garbern JC, Lee RT. Cardiac Stem Cell Therapy and the Promise of Heart Regeneration. *Cell Stem Cell*. 2013;12(6):689-698. doi:10.1016/j.stem.2013.05.008.
6. Thies RS, Murry CE. The advancement of human pluripotent stem cell-derived therapies into the clinic. *Development*. 2015;142(18):3077-3084. doi:10.1242/dev.126482.
7. Chong JJH, Yang X, Don CW, et al. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature*. 2014;510(7504):273-277. doi:10.1038/nature13233.
8. Kern S, Eichler H, Stoeve J, Klüter H, Bieback K. Comparative Analysis of Mesenchymal Stem Cells from Bone Marrow, Umbilical Cord Blood, or Adipose Tissue. *Stem Cells*. 2006;24(5):1294-1301. doi:10.1634/stemcells.2005-0342.
9. Wyles CC, Houdek MT, Behfar A, Sierra RJ. Mesenchymal stem cell therapy for osteoarthritis: current perspectives. *Stem Cells Cloning*. 2015;8:117-124. doi:10.2147/SCCAA.S68073.
10. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998;282(5391):1145-1147.
11. Murry CE, Keller G. Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons from Embryonic Development. *Cell*. 2008;132(4):661-680. doi:10.1016/J.CELL.2008.02.008.
12. Kokkinopoulos I, Ishida H, Saba R, Coppen S, Suzuki K, Yashiro K. Cardiomyocyte differentiation from mouse embryonic stem cells using a simple and defined protocol. *Dev Dyn*. 2016;245(2):157-165. doi:10.1002/dvdy.24366.
13. Bai F, Ho Lim C, Jia J, et al. Directed Differentiation of Embryonic Stem Cells Into Cardiomyocytes by Bacterial Injection of Defined Transcription Factors. *Sci Rep*. 2015;5:15014. doi:10.1038/srep15014.
14. Laflamme MA, Chen KY, Naumova A V, et al. Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat Biotechnol*. 2007;25(9):1015-1024. doi:10.1038/nbt1327.
15. Graichen R, Xu X, Braam SR, et al. Enhanced cardiomyogenesis of human embryonic stem cells by a small molecular inhibitor of p38 MAPK. *Differentiation*. 2008;76(4):357-370. doi:10.1111/j.1432-0436.2007.00236.x.
16. Lee S-H, Lumelsky N, Studer L, Auerbach JM, McKay RD. Efficient generation of midbrain and hindbrain neurons from mouse embryonic stem cells. *Nat Biotechnol*. 2000;18(6):675-679. doi:10.1038/76536.
17. Perrier AL, Tabar V, Barberi T, et al. Derivation of midbrain dopamine neurons from human embryonic stem cells. *Proc Natl Acad Sci U S A*. 2004;101(34):12543-12548. doi:10.1073/pnas.0404700101.
18. Srijaya TC, Pradeep PJ, Zain RB, et al. The Promise of Human Induced Pluripotent Stem Cells in Dental Research. *Stem Cells Int*. 2012;10. doi:10.1155/2012/423868.
19. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006;126(4):663-676. doi:10.1016/j.cell.2006.07.024.
20. Singh VK, Kalsan M, Kumar N, Saini A, Chandra R. Induced pluripotent stem

cells: applications in regenerative medicine, disease modeling, and drug discovery. *Front cell Dev Biol*. 2015;3:2. doi:10.3389/fcell.2015.00002.
21. Mandai M, Watanabe A, Kurimoto Y, et al. Autologous Induced Stem-Cell–Derived Retinal Cells for Macular Degeneration. *N Engl J Med*. 2017;376(11):1038-1046. doi:10.1056/NEJMoa1608368.
22. Kuriyan AE, Albini TA, Townsend JH, et al. Vision Loss after Intravitreal Injection of Autologous “Stem Cells” for AMD. *N Engl J Med*. 2017;376(11):1047-1053. doi:10.1056/NEJMoa1609583.
23. Li F, Hu J, He T-C. iPSC-based treatment of age-related macular degeneration (AMD): The path to success requires more than blind faith. *Genes Dis*. 2017;4(2):41-42. doi:10.1016/J.GENDIS.2017.03.001.
24. Dimos JT, Rodolfa KT, Niakan KK, et al. Induced Pluripotent Stem Cells Generated from Patients with ALS Can Be Differentiated into Motor Neurons. *Science*. 2008;321(5893):1218-1221. doi:10.1126/science.1158799.
25. Itzhaki I, Maizels L, Huber I, et al. Modelling the long QT syndrome with induced pluripotent stem cells. *Nature*. 2011;471(7337):225-229. doi:10.1038/nature09747.
26. Ebert AD, Liang P, Wu JC. Induced pluripotent stem cells as a disease modeling and drug screening platform. *J Cardiovasc Pharmacol*. 2012;60(4):408-416. doi:10.1097/FJC.0b013e318247f642.
27. Gutierrez-Aranda I, Ramos-Mejia V, Bueno C, et al. Human induced pluripotent stem cells develop teratoma more efficiently and faster than human embryonic stem cells regardless the site of injection. *Stem Cells*. 2010;28(9):1568-1570. doi:10.1002/stem.471.
28. Lensch MW, Schlaeger TM, Zon LI, Daley GQ. Teratoma Formation Assays with Human Embryonic Stem Cells: A Rationale for One Type of Human-Animal Chimera. *Cell Stem Cell*. 2007;1(3):253-258. doi:10.1016/j.stem.2007.07.019.
29. Fong C-Y, Gauthaman K, Bongso A. Teratomas from pluripotent stem cells: A clinical hurdle. *J Cell Biochem*. 2010;111(4):769-781. doi:10.1002/jcb.22775.
30. Cao F, Lin S, Xie X, et al. In Vivo Visualization of Embryonic Stem Cell Survival, Proliferation, and Migration After Cardiac Delivery. *Circulation*. 2006;113(7):1005-1014. doi:10.1161/CIRCULATIONAHA.105.588954.
31. Schuldiner M, Eiges R, Eden A, et al. Induced neuronal differentiation of human embryonic stem cells. *Brain Res*. 2001;913(2):201-205.
32. Fujikawa T, Oh S-H, Pi L, Hatch HM, Shupe T, Petersen BE. Teratoma Formation Leads to Failure of Treatment for Type I Diabetes Using Embryonic Stem Cell-Derived Insulin-Producing Cells. *Am J Pathol*. 2005;166(6):1781-1791. doi:10.1016/S0002-9440(10)62488-1.
33. Medvedev SP, Shevchenko AI, Zakian SM. Induced Pluripotent Stem Cells: Problems and Advantages when Applying them in Regenerative Medicine. *Acta Naturae*. 2010;2(2):18-28.
34. Takaoka K, Hamada H. Cell fate decisions and axis determination in the early mouse embryo. *Development*. 2012;139(1):3-14. doi:10.1242/dev.060095.
35. Johnson MH, Maro B, Takeichi M. The role of cell adhesion in the synchronization and orientation of polarization in 8-cell mouse blastomeres. *J Embryol Exp Morphol*. 1986;93:239-255.
36. Johnson MH, Ziomek CA. The foundation of two distinct cell lineages within the mouse morula. *Cell*. 1981;24(1):71-80.
37. Pan D. Hippo signaling in organ size control. *Genes Dev*. 2007;21(8):886-897. doi:10.1101/gad.1536007.
38. Nishioka N, Inoue K, Adachi K, et al. The Hippo signaling pathway components Lats and Yap pattern Tead4 activity to distinguish mouse trophectoderm from inner cell mass. *Dev Cell*. 2009;16(3):398-410. doi:10.1016/j.devcel.2009.02.003.
39. Nichols J, Zevnik B, Anastassiadis K, et al. Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell*. 1998;95(3):379-391.

40. Avilion AA, Nicolis SK, Pevny LH, Perez L, Vivian N, Lovell-Badge R. Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev*. 2003;17(1):126-140. doi:10.1101/gad.224503.
41. Mitsui K, Tokuzawa Y, Itoh H, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*. 2003;113(5):631-642.
42. Russ AP, Wattler S, Colledge WH, et al. Eomesodermin is required for mouse trophoblast development and mesoderm formation. *Nature*. 2000;404(6773):95-99. doi:10.1038/35003601.
43. Ralston A, Cox BJ, Nishioka N, et al. Gata3 regulates trophoblast development downstream of Tead4 and in parallel to Cdx2. *Development*. 2010;137(3):395-403. doi:10.1242/dev.038828.
44. Niwa H, Toyooka Y, Shimosato D, et al. Interaction between Oct3/4 and Cdx2 Determines Trophectoderm Differentiation. *Cell*. 2005;123(5):917-929. doi:10.1016/j.cell.2005.08.040.
45. Barcroft LC, Offenberg H, Thomsen P, Watson AJ. Aquaporin proteins in murine trophectoderm mediate transepithelial water movements during cavitation. *Dev Biol*. 2003;256(2):342-354.
46. Plusa B, Piliszek A, Frankenberg S, Artus J, Hadjantonakis A-K. Distinct sequential cell behaviours direct primitive endoderm formation in the mouse blastocyst. *Development*. 2008;135(18):3081-3091. doi:10.1242/dev.021519.
47. Morrisey EE, Tang Z, Sigrist K, et al. GATA6 regulates HNF4 and is required for differentiation of visceral endoderm in the mouse embryo. *Genes Dev*. 1998;12(22):3579-3590.
48. Koutsourakis M, Langeveld A, Patient R, Beddington R, Grosveld F. The transcription factor GATA6 is essential for early extraembryonic development. *Development*. 1999;126(9):723-732.
49. Niakan KK, Ji H, Maehr R, et al. Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev*. 2010;24(3):312-326. doi:10.1101/gad.1833510.
50. Artus J, Piliszek A, Hadjantonakis A-K. The primitive endoderm lineage of the mouse blastocyst: Sequential transcription factor activation and regulation of differentiation by Sox17. *Dev Biol*. 2011;350(2):393-404. doi:10.1016/j.ydbio.2010.12.007.
51. Frankenberg S, Gerbe F, Bessonnard S, et al. Primitive Endoderm Differentiates via a Three-Step Mechanism Involving Nanog and RTK Signaling. *Dev Cell*. 2011;21(6):1005-1013. doi:10.1016/j.devcel.2011.10.019.
52. Abe K, Niwa H, Iwase K, et al. Endoderm-Specific Gene Expression in Embryonic Stem Cells Differentiated to Embryoid Bodies. *Exp Cell Res*. 1996;229(1):27-34. doi:10.1006/excr.1996.0340.
53. Bedzhov I, Graham SJL, Leung CY, Zernicka-Goetz M. Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. *Philos Trans R Soc B Biol Sci*. 2014;369(1657):20130538-20130538. doi:10.1098/rstb.2013.0538.
54. Xiao L-J, Chang H, Ding N-Z, Ni H, Kadomatsu K, Yang Z-M. Basigin expression and hormonal regulation in mouse uterus during the peri-implantation period. *Mol Reprod Dev*. 2002;63(1):47-54. doi:10.1002/mrd.10128.
55. Basak S, Dhar R, Das C. Steroids modulate the expression of alpha4 integrin in mouse blastocysts and uterus during implantation. *Biol Reprod*. 2002;66(6):1784-1789.
56. Vinketova K, Mourdjeva M, Oreshkova T. Human Decidual Stromal Cells as a Component of the Implantation Niche and a Modulator of Maternal Immunity. *J Pregnancy*. 2016;2016:8689436. doi:10.1155/2016/8689436.
57. Nancy P, Tagliani E, Tay C-S, Asp P, Levy DE, Erlebacher A. Chemokine Gene Silencing in Decidual Stromal Cells Limits T Cell Access to the Maternal-Fetal Interface. *Science*. 2012;336(6086):1317-1321. doi:10.1126/science.1220030.

58. Donnison M, Beaton A, Davey HW, Broadhurst R, L'Huillier P, Pfeffer PL. Loss of the extraembryonic ectoderm in Elf5 mutants leads to defects in embryonic patterning. *Development*. 2005;132(10):2299-2308. doi:10.1242/dev.01819.
59. Stemple DL. Vertebrate development: the subtle art of germ-layer specification. *Curr Biol*. 2001;11(21):878-81.
60. Williams M, Burdsal C, Periasamy A, Lewandoski M, Sutherland A. Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev Dyn*. 2012;241(2):270-283. doi:10.1002/dvdy.23711.
61. Solnica-Krezel L, Sepich DS. Gastrulation: Making and Shaping Germ Layers. *Annu Rev Cell Dev Biol*. 2012;28(1):687-717. doi:10.1146/annurev-cellbio-092910-154043.
62. Tam PPL, Loebel DAF. Gene function in mouse embryogenesis: get set for gastrulation. *Nat Rev Genet*. 2007;8(5):368-381. doi:10.1038/nrg2084.
63. Carlson BM, Carlson BM. Formation of Germ Layers and Early Derivatives. In: *Human Embryology and Developmental Biology*. Elsevier; 2014:75-91. doi:10.1016/B978-1-4557-2794-0.00005-X.
64. Kwon GS, Viotti M, Hadjantonakis A-K. The endoderm of the mouse embryo arises by dynamic widespread intercalation of embryonic and extraembryonic lineages. *Dev Cell*. 2008;15(4):509-520. doi:10.1016/j.devcel.2008.07.017.
65. Lawson KA, Meneses JJ, Pedersen RA. Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development*. 1991;113(3):891-911.
66. Tada S, Era T, Furusawa C, et al. Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development*. 2005;132(19):4363-4374. doi:10.1242/dev.02005.
67. Leptin M. Review Gastrulation Movements: the Logic and the Nuts and Bolts. *Dev Cell*. 2005;8:305-320. doi:10.1016/j.devcel.2005.02.007.
68. Murry CE, Keller G. Differentiation of Embryonic Stem Cells to Clinically Relevant Populations: Lessons from Embryonic Development. *Cell*. 2008;132(4):661-680. doi:10.1016/j.cell.2008.02.008.
69. Simon MC, Keith B. The role of oxygen availability in embryonic development and stem cell function. *Nat Rev Mol Cell Biol*. 2008;9(4):285-296. doi:10.1038/nrm2354.
70. Okazaki K, Maltepe E. Oxygen, epigenetics and stem cell fate. *Regen Med*. 2006;1(1):71-83. doi:10.2217/17460751.1.1.71.
71. Fischer B, Bavister BD. Oxygen tension in the oviduct and uterus of rhesus monkeys, hamsters and rabbits. *J Reprod Fertil*. 1993;99(2):673-679.
72. Wang Y, Zhao S. *Vascular Biology of the Placenta*. Morgan & Claypool Life Sciences; 2010.
73. Lee YM, Jeong C-H, Koo S-Y, et al. Determination of hypoxic region by hypoxia marker in developing mouse embryos in vivo: A possible signal for vessel development. *Dev Dyn*. 2001;220(2):175-186. doi:10.1002/1097-0177(20010201)220:2<175::AID-DVDY1101>3.0.CO;2-F.
74. Pringle KG, Kind KL, Thompson JG, Roberts CT. Complex Interactions Between Hypoxia Inducible Factors, Insulin-Like Growth Factor-II and Oxygen in Early Murine Trophoblasts. *Placenta*. 2007;28(11-12):1147-1157. doi:10.1016/j.placenta.2007.05.009.
75. Gross MW, Karbach U, Groebe K, Franko AJ, Mueller-Klieser W. Calibration of misonidazole labeling by simultaneous measurement of oxygen tension and labeling density in multicellular spheroids. *Int J cancer*. 1995;61(4):567-573.
76. Dunwoodie SL. The role of hypoxia in development of the Mammalian embryo. *Dev Cell*. 2009;17(6):755-773. doi:10.1016/j.devcel.2009.11.008.
77. Olive PL, Vikse C, Trotter MJ. Measurement of oxygen diffusion distance in tumor cubes using a fluorescent hypoxia probe. *Int J Radiat Oncol Biol Phys*. 1992;22(3):397-402.
78. Krogh A. The number and distribution of capillaries in muscles with calculations

of the oxygen pressure head necessary for supplying the tissue. *J Physiol*. 1919;52(6):409-415.
79. Gatenby RA, Gillies RJ. Why do cancers have high aerobic glycolysis? *Nat Rev Cancer*. 2004;4(11):891-899. doi:10.1038/nrc1478.
80. Fisher S a, Burggren WW. Role of hypoxia in the evolution and development of the cardiovascular system. *Antioxid Redox Signal*. 2007;9(9):1339-1352. doi:10.1089/ars.2007.1704.
81. Mohyeldin A, Garzón-Muvdi T, Quiñones-Hinojosa A. Oxygen in Stem Cell Biology: A Critical Component of the Stem Cell Niche. *Cell Stem Cell*. 2010;7(2):150-161. doi:10.1016/j.stem.2010.07.007.
82. Turing AM. The Chemical Basis of Morphogenesis. *Philos Trans R Soc B Biol Sci*. 1952;237(641):37-72. doi:10.1098/rstb.1952.0012.
83. Tabata T, Takei Y. Morphogens, their identification and regulation. *Development*. 2004;131(4):703-712. doi:10.1242/dev.01043.
84. Rogers KW, Schier AF. Morphogen Gradients: From Generation to Interpretation. *Annu Rev Cell Dev Biol*. 2011;27(1):377-407. doi:10.1146/annurev-cellbio-092910-154148.
85. Mullor JL, Calleja M, Capdevila J, Guerrero I. Hedgehog activity, independent of decapentaplegic, participates in wing disc patterning. *Development*. 1997;124(6):1227-1237.
86. Zecca M, Basler K, Struhl G. Direct and long-range action of a wingless morphogen gradient. *Cell*. 1996;87(5):833-844.
87. Lecuit T, Cohen SM. Dpp receptor levels contribute to shaping the Dpp morphogen gradient in the Drosophila wing imaginal disc. *Development*. 1998;125(24):4901-4907.
88. Chen Y, Schier AF. The zebrafish Nodal signal Squint functions as a morphogen. *Nature*. 2001;411(6837):607-610. doi:10.1038/35079121.
89. Lopez-Barneo J, Pardal R, Ortega-Sáenz P. Cellular mechanism of oxygen sensing. *Annu Rev Physiol*. 2001;63:259-287. doi:10.1146/annurev.physiol.63.1.259.
90. Jones RD, Hancock JT, Morice AH. NADPH oxidase: a universal oxygen sensor? *Free Radic Biol Med*. 2000;29(5):416-424.
91. Guzy RD, Hoyos B, Robin E, et al. Mitochondrial complex III is required for hypoxia-induced ROS production and cellular oxygen sensing. *Cell Metab*. 2005;1(6):401-408. doi:10.1016/j.cmet.2005.05.001.
92. Archer SL, Wu X-C, Thébaud B, Moudgil R, Hashimoto K, Michelakis ED. O2 sensing in the human ductus arteriosus: redox-sensitive K+ channels are regulated by mitochondria-derived hydrogen peroxide. *Biol Chem*. 2004;385(3-4):205-216. doi:10.1515/BC.2004.014.
93. Giaccia AJ, Simon MC, Johnson R. The biology of hypoxia: the role of oxygen sensing in development, normal function, and disease. *Genes Dev*. 2004;18(18):2183-2194. doi:10.1101/gad.1243304.
94. Ke Q, Costa M. Hypoxia-Inducible Factor-1 (HIF-1). *Mol Pharmacol*. 2006;70(5):1469-1480. doi:10.1124/mol.106.027029.
95. Jiang BH, Semenza GL, Bauer C, Marti HH. Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O2 tension. *Am J Physiol Physiol*. 1996;271(4):C1172-C1180. doi:10.1152/ajpcell.1996.271.4.C1172.
96. Schofield CJ, Ratcliffe PJ. Oxygen sensing by HIF hydroxylases. *Nat Rev Mol Cell Biol*. 2004;5(5):343-354. doi:10.1038/nrm1366.
97. Schofield CJ, Zhang Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr Opin Struct Biol*. 1999;9(6):722-731.
98. Lando D, Peet DJ, Gorman JJ, Whelan DA, Whitelaw ML, Bruick RK. FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. *Genes Dev*. 2002;16(12):1466-1471. doi:10.1101/gad.991402.

99. Ramírez-Bergeron DL, Runge A, Adelman DM, Gohil M, Simon MC. HIF-Dependent Hematopoietic Factors Regulate the Development of the Embryonic Vasculature. *Dev Cell*. 2006;11(1):81-92. doi:10.1016/j.devcel.2006.04.018.
100. Maltepe E, Schmidt J V., Baunoch D, Bradfield CA, Simon MC. Abnormal angiogenesis and responses to glucose and oxygen deprivation in mice lacking the protein ARNT. *Nature*. 1997;386(6623):403-407. doi:10.1038/386403a0.
101. Iyer N V, Kotch LE, Agani F, et al. Cellular and developmental control of O2 homeostasis by hypoxia-inducible factor 1 alpha. *Genes Dev*. 1998;12(2):149-162.
102. Ryan HE, Lo J, Johnson RS. HIF-1alpha is required for solid tumor formation and embryonic vascularization. *EMBO J*. 1998;17(11):3005-3015. doi:10.1093/emboj/17.11.3005.
103. Yin Z, Haynie J, Yang X, et al. The essential role of Cited2, a negative regulator for HIF-1alpha, in heart development and neurulation. *Proc Natl Acad Sci U S A*. 2002;99(16):10488-10493. doi:10.1073/pnas.162371799.
104. Cowden Dahl KD, Fryer BH, Mack FA, et al. Hypoxia-Inducible Factors 1 and 2 Regulate Trophoblast Differentiation. *Mol Cell Biol*. 2005;25(23):10479-10491. doi:10.1128/MCB.25.23.10479-10491.2005.
105. Abbott BD, Buckalew AR. Placental defects in ARNT-knockout conceptus correlate with localized decreases in VEGF-R2, Ang-1, and Tie-2. *Dev Dyn*. 2000;219(4):526-538. doi:10.1002/1097-0177(2000)9999:9999<::AID-DVDY1080>3.0.CO;2-N.
106. Kozak KR, Abbott B, Hankinson O. ARNT-Deficient Mice and Placental Differentiation. *Dev Biol*. 1997;191(2):297-305. doi:10.1006/dbio.1997.8758.
107. Genbacev O, Zhou Y, Ludlow JW, Fisher SJ. Regulation of human placental development by oxygen tension. *Science*. 1997;277(5332):1669-1672.
108. Krishnan J, Ahuja P, Bodenmann S, et al. Essential Role of Developmentally Activated Hypoxia-Inducible Factor 1 for Cardiac Morphogenesis and Function. *Circ Res*. 2008;103(10):1139-1146. doi:10.1161/01.RES.0000338613.89841.c1.
109. Compernolle V, Brusselmans K, Franco D, et al. Cardia bifida, defective heart development and abnormal neural crest migration in embryos lacking hypoxia-inducible factor-1alpha. *Cardiovasc Res*. 2003;60(3):569-579.
110. Tian H, Hammer RE, Matsumoto AM, Russell DW, McKnight SL. The hypoxia-responsive transcription factor EPAS1 is essential for catecholamine homeostasis and protection against heart failure during embryonic development. *Genes Dev*. 1998;12(21):3320-3324.
111. Scortegagna M, Ding K, Oktay Y, et al. Multiple organ pathology, metabolic abnormalities and impaired homeostasis of reactive oxygen species in Epas1−/− mice. *Nat Genet*. 2003;35(4):331-340. doi:10.1038/ng1266.
112. Compernolle V, Brusselmans K, Acker T, et al. Erratum: Loss of HIF-2α and inhibition of VEGF impair fetal lung maturation, whereas treatment with VEGF prevents fatal respiratory distress in premature mice. *Nat Med*. 2002;8(7):702-710. doi:10.1038/nm721.
113. Covello KL, Kehler J, Yu H, et al. HIF-2alpha regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes Dev*. 2006;20(5):557-570. doi:10.1101/gad.1399906.
114. Forristal CE, Wright KL, Hanley NA, Oreffo ROC, Houghton FD. Hypoxia inducible factors regulate pluripotency and proliferation in human embryonic stem cells cultured at reduced oxygen tensions. *Reproduction*. 2010;139(1):85-97. doi:10.1530/REP-09-0300.
115. Petruzzelli R, Christensen DR, Parry KL, Sanchez-Elsner T, Houghton FD. HIF-2α Regulates NANOG Expression in Human Embryonic Stem Cells following Hypoxia and Reoxygenation through the Interaction with an Oct-Sox Cis Regulatory Element. Rocha S, ed. *PLoS One*. 2014;9(10):e108309. doi:10.1371/journal.pone.0108309.
116. Kranc KR, Oliveira D V, Armesilla-Diaz A, et al. Acute Loss of Cited2 Impairs Nanog Expression and Decreases Self-Renewal of Mouse Embryonic Stem

Cells. *Stem Cells*. 2015;33(3):699-712. doi:10.1002/stem.1889.
117. Li L, Candelario KM, Thomas K, et al. Hypoxia Inducible Factor-1 (HIF-1) Is Required for Neural Stem Cell Maintenance and Vascular Stability in the Adult Mouse SVZ. *J Neurosci*. 2014;34(50):16713-16719. doi:10.1523/JNEUROSCI.4590-13.2014.
118. Zhao Y, Matsuo-Takasaki M, Tsuboi I, et al. Dual Functions of Hypoxia-Inducible Factor 1 Alpha for the Commitment of Mouse Embryonic Stem Cells Toward a Neural Lineage. *Stem Cells Dev*. 2014;23(18):2143-2155. doi:10.1089/scd.2013.0278.
119. Ng K-M, Lee Y-K, Chan Y-C, et al. Exogenous expression of HIF-1 alpha promotes cardiac differentiation of embryonic stem cells. *J Mol Cell Cardiol*. 2010;48(6):1129-1137. doi:10.1016/j.yjmcc.2010.01.015.
120. Dengler VL, Galbraith M, Espinosa JM. Transcriptional regulation by hypoxia inducible factors. *Crit Rev Biochem Mol Biol*. 2014;49(1):1-15. doi:10.3109/10409238.2013.838205.
121. Rose NR, McDonough MA, King ONF, Kawamura A, Schofield CJ. Inhibition of 2-oxoglutarate dependent oxygenases. *Chem Soc Rev*. 2011;40(8):4364. doi:10.1039/c0cs00203h.
122. Martinez S, Hausinger RP. Catalytic Mechanisms of Fe(II)-and 2-Oxoglutarate-Dependent Oxygenases. doi:10.1074/jbc.R115.648691.
123. Kawai Y, Ono E, Mizutani M. Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. *Plant J*. 2014;78(2):328-343. doi:10.1111/tpj.12479.
124. Salminen A, Kauppinen A, Kaarniranta K. 2-Oxoglutarate-dependent dioxygenases are sensors of energy metabolism, oxygen availability, and iron homeostasis: potential role in the regulation of aging process. *Cell Mol Life Sci*. 2015;72(20):3897-3914. doi:10.1007/s00018-015-1978-z.
125. Kuiper C, Vissers MCM. Ascorbate as a co-factor for fe- and 2-oxoglutarate dependent dioxygenases: physiological activity in tumor growth and progression. *Front Oncol*. 2014;4:359. doi:10.3389/fonc.2014.00359.
126. Flashman E, Davies SL, Yeoh KK, Schofield CJ. Investigating the dependence of the hypoxia-inducible factor hydroxylases (factor inhibiting HIF and prolyl hydroxylase domain 2) on ascorbate and other reducing agents. *Biochem J*. 2010;427(1):135-142. doi:10.1042/BJ20091609.
127. Markolovic S, Wilkins SE, Schofield CJ. Protein Hydroxylation Catalyzed by 2-Oxoglutarate-dependent Oxygenases. *J Biol Chem*. 2015;290(34):20712-20722. doi:10.1074/jbc.R115.662627.
128. Klose RJ, Kallin EM, Zhang Y. JmjC-domain-containing proteins and histone demethylation. *Nat Rev Genet*. 2006;7(9):715-727. doi:10.1038/nrg1945.
129. Ezashi T, Das P, Roberts RM. Low O2 tensions and the prevention of differentiation of hES cells. *Proc Natl Acad Sci U S A*. 2005;102(13):4783-4788. doi:10.1073/pnas.0501283102.
130. Prasad SM, Czepiel M, Cetinkaya C, et al. Continuous hypoxic culturing maintains activation of Notch and allows long-term propagation of human embryonic stem cells without spontaneous differentiation. *Cell Prolif*. 2009;42(1):63-74. doi:10.1111/j.1365-2184.2008.00571.x.
131. Lengner CJ, Gimelbrant A a, Erwin J a, et al. Derivation of pre-X inactivation human embryonic stem cells under physiological oxygen concentrations. *Cell*. 2010;141(5):872-883. doi:10.1016/j.cell.2010.04.010.
132. Forsyth NR, Kay A, Hampson K, Downing A, Talbot R, McWhir J. Transcriptome alterations due to physiological normoxic (2% O_2) culture of human embryonic stem cells. *Regen Med*. 2008;3(6):817-833. doi:10.2217/17460751.3.6.817.
133. Chen H-F, Kuo H-C, Chen W, Wu F-C, Yang Y-S, Ho H-N. A reduced oxygen tension (5%) is not beneficial for maintaining human embryonic stem cells in the undifferentiated state with short splitting intervals. doi:10.1093/humrep/den345.
134. Nä Rvä E, Pursiheimo J-P, Laiho A, et al. Continuous Hypoxic Culturing of Human Embryonic Stem Cells Enhances SSEA-3 and MYC Levels. *PLoS One*.

2013;8(11). doi:10.1371/journal.pone.0078847.
135. Rinaudo P, Giritharan G, Talbi S, Dobson A, Schultz R. Effects of oxygen tension on gene expression in preimplantation mouse embryos. *Fertil Steril*. 2006;86(4):1265.e1-1265.e36. doi:10.1016/j.fertnstert.2006.05.017.
136. Orsi NM, Leese HJ. Protection against reactive oxygen species during mouse preimplantation embryo development: Role of EDTA, oxygen tension, catalase, superoxide dismutase and pyruvate. *Mol Reprod Dev*. 2001;59(1):44-53. doi:10.1002/mrd.1006.
137. Jeong C-H, Lee H-J, Cha J-H, et al. Hypoxia-inducible Factor-1 Inhibits Self-renewal of Mouse Embryonic Stem Cells in Vitro via Negative Regulation of the Leukemia Inhibitory Factor-STAT3 Pathway. 2007. doi:10.1074/jbc.M700534200.
138. Pimton P, Lecht S, Stabler CT, Johannes G, Schulman ES, Lelkes PI. Hypoxia enhances differentiation of mouse embryonic stem cells into definitive endoderm and distal lung cells. *Stem Cells Dev*. 2015;24(5):663-676. doi:10.1089/scd.2014.0343.
139. Katsuda T, Teratani T, Chowdhury MM, Ochiya T, Sakai Y. Hypoxia efficiently induces differentiation of mouse embryonic stem cells into endodermal and hepatic progenitor cells. *Biochem Eng J*. 2013;74:95-101. doi:10.1016/j.bej.2013.02.012.
140. Williams K, Christensen J, Helin K. DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep*. 2012;13(1):28-35. doi:10.1038/embor.2011.233.
141. Goldberg AD, Allis CD, Bernstein E. Epigenetics: A Landscape Takes Shape. *Cell*. 2007;128(4):635-638. doi:10.1016/j.cell.2007.02.006.
142. Tollervey JR, Lunyak V V. Epigenetics: judge, jury and executioner of stem cell fate. *Epigenetics*. 2012;7(8):823-840. doi:10.4161/epi.21141.
143. Handy DE, Castro R, Loscalzo J. Epigenetic modifications: basic mechanisms and role in cardiovascular disease. *Circulation*. 2011;123(19):2145-2156. doi:10.1161/CIRCULATIONAHA.110.956839.
144. Rajender S, Avery K, Agarwal A. Epigenetics, spermatogenesis and male infertility. *Mutat Res Mutat Res*. 2011;727(3):62-71. doi:10.1016/j.mrrev.2011.04.002.
145. Molina-Serrano D, Schiza V, Kirmizis A. Cross-talk among epigenetic modifications: lessons from histone arginine methylation. *Biochem Soc Trans*. 2013;41(3):751-759. doi:10.1042/BST20130003.
146. Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell*. 2013;154(3):490-503. doi:10.1016/j.cell.2013.07.011.
147. Varga-Weisz P. ATP-dependent chromatin remodeling factors: Nucleosome shufflers with many missions. *Oncogene*. 2001;20(24):3076-3085. doi:10.1038/sj.onc.1204332.
148. Hargreaves DC, Crabtree GR. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res*. 2011;21(3):396-420. doi:10.1038/cr.2011.32.
149. Greco CM, Condorelli G. Epigenetic modifications and noncoding RNAs in cardiac hypertrophy and failure. *Nat Rev Cardiol*. 2015;12(8):488-497. doi:10.1038/nrcardio.2015.71.
150. Kaikkonen MU, Lam MTY, Glass CK. Non-coding RNAs as regulators of gene expression and epigenetics. *Cardiovasc Res*. 2011;90(3):430-440. doi:10.1093/cvr/cvr097.
151. Chen J, Xue Y. Emerging roles of non-coding RNAs in epigenetic regulation. *Sci China Life Sci*. 2016;59(3):227-235. doi:10.1007/s11427-016-5010-0.
152. Peschansky VJ, Wahlestedt C. Non-coding RNAs as direct and indirect modulators of epigenetic regulation. *Epigenetics*. 2014;9(1):3-12. doi:10.4161/epi.27473.
153. Strahl BD, Allis CD. The language of covalent histone modifications. *Nature*. 2000;403(6765):41-45. doi:10.1038/47412.

154. Eberharter A, Becker PB. Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. *EMBO Rep*. 2002;3(3):224-229. doi:10.1093/embo-reports/kvf053.
155. Rossetto D, Avvakumov N, Côté J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics*. 2012;7(10):1098-1108. doi:10.4161/epi.21975.
156. Kouzarides T. Chromatin Modifications and Their Function. *Cell*. 2007;128(4):693-705. doi:10.1016/j.cell.2007.02.005.
157. Hitchler MJ, Domann FE. An epigenetic perspective on the free radical theory of development. *Free Radic Biol Med*. 2007;43(7):1023-1036. doi:10.1016/j.freeradbiomed.2007.06.027.
158. Saxonov S, Berg P, Brutlag DL. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci*. 2006;103(5):1412-1417. doi:10.1073/pnas.0510310103.
159. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99(3):247-257.
160. Nan X, Ng H-H, Johnson CA, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*. 1998;393(6683):386-389. doi:10.1038/30764.
161. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol*. 2010;11(9):607-620. doi:10.1038/nrm2950.
162. Morgan HD, Dean W, Coker HA, Reik W, Petersen-Mahrt SK. Activation-induced Cytidine Deaminase Deaminates 5-Methylcytosine in DNA and Is Expressed in Pluripotent Tissues. *J Biol Chem*. 2004;279(50):52353-52360. doi:10.1074/jbc.M407695200.
163. Zhu J-K. Active DNA Demethylation Mediated by DNA Glycosylases. *Annu Rev Genet*. 2009;43(1):143-166. doi:10.1146/annurev-genet-102108-134205.
164. Barreto G, Schäfer A, Marhold J, et al. Gadd45a promotes epigenetic gene activation by repair-mediated DNA demethylation. *Nature*. 2007;445(7128):671-675. doi:10.1038/nature05515.
165. Métivier R, Gallais R, Tiffoche C, et al. Cyclical DNA methylation of a transcriptionally active promoter. *Nature*. 2008;452(7183):45-50. doi:10.1038/nature06544.
166. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. *Nature*. 2013;502(7472):472-479. doi:10.1038/nature12750.
167. Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R. The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J*. 1972;126(4):781-790.
168. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*. 2009;324(5929):930-935. doi:10.1126/science.1170116.
169. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science*. 2009;324(5929):929-930. doi:10.1126/science.1169786.
170. Globisch D, Münzel M, Müller M, et al. Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One*. 2010;5(12):e15367. doi:10.1371/journal.pone.0015367.
171. Lorsbach RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR. TET1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). *Leukemia*. 2003;17(3):637-641. doi:10.1038/sj.leu.2402834.
172. Iyer LM, Tahiliani M, Rao A, Aravind L. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle*. 2009;8(11):1698-1710. doi:10.4161/cc.8.11.8580.
173. Tan L, Shi YG. Tet family proteins and 5-hydroxymethylcytosine in development and disease. *Development*. 2012;139(11):1895-1902. doi:10.1242/dev.070771.

174. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res*. 2010;20(12):1390-1393. doi:10.1038/cr.2010.156.
175. Jin S-G, Zhang Z-M, Dunwell TL, et al. Tet3 Reads 5-Carboxylcytosine through Its CXXC Domain and Is a Potential Guardian against Neurodegeneration. *Cell Rep*. 2016;14(3):493-505. doi:10.1016/j.celrep.2015.12.044.
176. Xu Y, Xu C, Kato A, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell*. 2012;151(6):1200-1213. doi:10.1016/j.cell.2012.11.014.
177. Ko M, An J, Bandukwala HS, et al. Modulation of TET2 expression and 5-methylcytosine oxidation by the CXXC domain protein IDAX. *Nature*. 2013;497(7447):122-126. doi:10.1038/nature12052.
178. Bachman M, Uribe-Lewis S, Yang X, Williams M, Murrell A, Balasubramanian S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat Chem*. 2014;6(12):1049-1055. doi:10.1038/nchem.2064.
179. Hahn MA, Szabó PE, Pfeifer GP. 5-Hydroxymethylcytosine: A stable or transient DNA modification? *Genomics*. 2014;104(5):314-323. doi:10.1016/j.ygeno.2014.08.015.
180. Münzel M, Globisch D, Carell T. 5-Hydroxymethylcytosine, the Sixth Base of the Genome. *Angew Chemie Int Ed*. 2011;50(29):6460-6468. doi:10.1002/anie.201101547.
181. Pastor WA, Aravind L, Rao A. TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat Publ Gr*. 2013;14. doi:10.1038/nrm3589.
182. Song C-X, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotechnol*. 2011;29(1):68-72. doi:10.1038/nbt.1732.
183. Xu Y, Wu F, Tan L, et al. Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Mol Cell*. 2011;42(4):451-464. doi:10.1016/j.molcel.2011.04.005.
184. Pastor WA, Pape UJ, Huang Y, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*. 2011;473(7347):394-397. doi:10.1038/nature10102.
185. Szulwach KE, Li X, Li Y, et al. Integrating 5-Hydroxymethylcytosine into the Epigenomic Landscape of Human Embryonic Stem Cells. Pearson CE, ed. *PLoS Genet*. 2011;7(6):e1002154. doi:10.1371/journal.pgen.1002154.
186. Ficz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*. 2011;473(7347):398-402. doi:10.1038/nature10008.
187. Lercher L, McDonough MA, El-Sagheer AH, et al. Structural insights into how 5-hydroxymethylation influences transcription factor binding. *Chem Commun*. 2014;50(15):1794-1796. doi:10.1039/c3cc48151d.
188. Valinluck V, Tsai H-H, Rogstad DK, Burdzy A, Bird A, Sowers LC. Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res*. 2004;32(14):4100-4108. doi:10.1093/nar/gkh739.
189. Li D, Guo B, Wu H, Tan L, Lu Q. TET Family of Dioxygenases: Crucial Roles and Underlying Mechanisms. *Cytogenet Genome Res*. 2015;146(3):171-180. doi:10.1159/000438853.
190. Vella P, Scelfo A, Jammula S, et al. Tet Proteins Connect the O-Linked N-acetylglucosamine Transferase Ogt to Chromatin in Embryonic Stem Cells. *Mol Cell*. 2013;49(4):645-656. doi:10.1016/j.molcel.2012.12.019.
191. Deplus R, Delatte B, Schwinn MK, et al. TET2 and TET3 regulate GlcNAcylation and H3K4 methylation through OGT and SET1/COMPASS. *EMBO J*. 2013;32357:645-655. doi:10.1038/emboj.2012.357.
192. Capotosti F, Guernier S, Lammers F, et al. O-GlcNAc Transferase Catalyzes

Site-Specific Proteolysis of HCF-1. *Cell*. 2011;144(3):376-388. doi:10.1016/j.cell.2010.12.030.
193. Voigt P, Tee W-W, Reinberg D. A double take on bivalent promoters. *Genes Dev*. 2013;27(12):1318-1338. doi:10.1101/gad.219626.113.
194. Wu H, D'Alessio AC, Ito S, et al. Dual functions of Tet1 in transcriptional regulation in mouse embryonic stem cells. *Nature*. 2011;473(7347):389-393. doi:10.1038/nature09934.
195. Williams K, Christensen J, Pedersen MT, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*. 2011;473(7347):343-348. doi:10.1038/nature10066.
196. Grzenda A, Lomberk G, Zhang J-S, Urrutia R. Sin3: Master scaffold and transcriptional corepressor. *Biochim Biophys Acta - Gene Regul Mech*. 2009;1789(6-8):443-450. doi:10.1016/j.bbagrm.2009.05.007.
197. He Y-F, Li B-Z, Li Z, et al. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. doi:10.1126/science.1210944.
198. Ito S, Shen L, Dai Q, et al. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science*. 2011;333(6047):1300-1303. doi:10.1126/science.1210597.
199. Pfaffeneder T, Hackner B, Truss M, et al. The discovery of 5-formylcytosine in embryonic stem cell DNA. *Angew Chem Int Ed Engl*. 2011;50(31):7008-7012. doi:10.1002/anie.201103899.
200. Fromme JC, Verdine GL. Base Excision Repair. In: *Advances in Protein Chemistry*. Vol 69. ; 2004:1-41. doi:10.1016/S0065-3233(04)69001-2.
201. Song C-X, He C. Potential functional roles of DNA demethylation intermediates. *Trends Biochem Sci*. 2013;38(10):480-484. doi:10.1016/j.tibs.2013.07.003.
202. He Y-F, Li B-Z, Li Z, et al. Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science*. 2011;333(6047):1303-1307. doi:10.1126/science.1210944.
203. Shen L, Wu H, Diep D, et al. Genome-wide Analysis Reveals TET- and TDG-Dependent 5-Methylcytosine Oxidation Dynamics. *Cell*. 2013;153(3):692-706. doi:10.1016/j.cell.2013.04.002.
204. Song C-X, Szulwach KE, Dai Q, et al. Genome-wide Profiling of 5-Formylcytosine Reveals Its Roles in Epigenetic Priming. *Cell*. 2013;153(3):678-691. doi:10.1016/j.cell.2013.04.001.
205. Neri F, Incarnato D, Krepelova A, et al. Single-Base Resolution Analysis of 5-Formyl and 5-Carboxyl Cytosine Reveals Promoter DNA Methylation Dynamics. *Cell Rep*. 2015;10(5):674-683. doi:10.1016/j.celrep.2015.01.008.
206. Hu L, Lu J, Cheng J, et al. Structural insight into substrate preference for TET-mediated oxidation. *Nature*. 2015;527(7576):118-122. doi:10.1038/nature15713.
207. An J, Rao A, Ko M. TET family dioxygenases and DNA demethylation in stem cells and cancers. *Exp Mol Med*. 2017;49(4):e323-e323. doi:10.1038/emm.2017.5.
208. Valinluck V, Sowers LC. Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1. *Cancer Res*. 2007;67(3):946-950. doi:10.1158/0008-5472.CAN-06-3123.
209. Ji D, Lin K, Song J, Wang Y. Effects of Tet-induced oxidation products of 5-methylcytosine on Dnmt1- and DNMT3a-mediated cytosine methylation. *Mol Biosyst*. 2014;10(7):1749. doi:10.1039/c4mb00150h.
210. Hashimoto H, Liu Y, Upadhyay AK, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res*. 2012;40(11):4841-4849. doi:10.1093/nar/gks155.
211. Torres-Padilla ME. Cell identity in the preimplantation mammalian embryo: an epigenetic perspective from the mouse. *Hum Reprod*. 2008;23(6):1246-1252. doi:10.1093/humrep/dem434.
212. Pontier DB, Gribnau J. Xist regulation and function explored. *Hum Genet*. 2011;130(2):223-236. doi:10.1007/s00439-011-1008-7.
213. Beard C, Li E, Jaenisch R. Loss of methylation activates Xist in somatic but not

in embryonic cells. *Genes Dev*. 1995;9(19):2325-2334. doi:10.1101/GAD.9.19.2325.
214. Reik W, Dean W, Walter J. Epigenetic Reprogramming in Mammalian Development. *Science*. 2001;293(5532):1089-1093. doi:10.1126/science.1063443.
215. Weber M, Hellmann I, Stadler MB, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*. 2007;39(4):457-466. doi:10.1038/ng1990.
216. Mohn F, Weber M, Rebhan M, et al. Lineage-Specific Polycomb Targets and De Novo DNA Methylation Define Restriction and Potential of Neuronal Progenitors. *Mol Cell*. 2008;30(6):755-766. doi:10.1016/j.molcel.2008.05.007.
217. Illingworth R, Kerr A, DeSousa D, et al. A Novel CpG Island Set Identifies Tissue-Specific Methylation at Developmental Gene Loci. Liu ET, ed. *PLoS Biol*. 2008;6(1):e22. doi:10.1371/journal.pbio.0060022.
218. Illingworth RS, Bird AP. CpG islands - "A rough guide." *FEBS Lett*. 2009;583(11):1713-1720. doi:10.1016/j.febslet.2009.04.012.
219. Wang L, Liu Z, Lin H, Ma D, Tao Q, Liu F. Epigenetic regulation of left–right asymmetry by DNA methylation. *EMBO J*. 2017;36(20):2987-2997. doi:10.15252/embj.201796580.
220. Kota SK, Feil R. Epigenetic Transitions in Germ Cell Development and Meiosis. *Dev Cell*. 2010;19(5):675-686. doi:10.1016/j.devcel.2010.10.009.
221. Canovas S, Ross PJ. Epigenetics in preimplantation mammalian development. *Theriogenology*. 2016;86(1):69-79. doi:10.1016/j.theriogenology.2016.04.020.
222. Lei H, Oh SP, Okano M, et al. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development*. 1996;122(10):3195-3205.
223. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell*. 1999;99(3):247-257.
224. Yin L-J, Zhang Y, Lv P-P, et al. Insufficient maintenance DNA methylation is associated with abnormal embryonic development. *BMC Med*. 2012;10(1):26. doi:10.1186/1741-7015-10-26.
225. Peters J. The role of genomic imprinting in biology and disease: an expanding view. *Nat Rev Genet*. 2014;15(8):517-530. doi:10.1038/nrg3766.
226. McEwen KR, Ferguson-Smith AC. Distinguishing epigenetic marks of developmental and imprinting regulation. *Epigenetics Chromatin*. 2010;3(1):2. doi:10.1186/1756-8935-3-2.
227. Dahl JA, Reiner AH, Klungland A, Wakayama T, Collas P. Histone H3 Lysine 27 Methylation Asymmetry on Developmentally-Regulated Promoters Distinguish the First Two Lineages in Mouse Preimplantation Embryos. Feil R, ed. *PLoS One*. 2010;5(2):e9150. doi:10.1371/journal.pone.0009150.
228. Burton A, Torres-Padilla M-E. Chromatin dynamics in the regulation of cell fate allocation during early embryogenesis. *Nat Rev Mol Cell Biol*. 2014;15(11):723-735. doi:10.1038/nrm3885.
229. Rugg-Gunn PJ, Cox BJ, Ralston A, Rossant J. Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc Natl Acad Sci U S A*. 2010;107(24):10783-10790. doi:10.1073/pnas.0914507107.
230. Torres-Padilla M-E, Parfitt D-E, Kouzarides T, Zernicka-Goetz M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature*. 2007;445(7124):214-218. doi:10.1038/nature05458.
231. Delhommeau F, Dupont S, Valle V Della, et al. Mutation in TET2 in Myeloid Cancers. *N Engl J Med*. 2009;360(22):2289-2301. doi:10.1056/NEJMoa0810069.
232. Ko M, Bandukwala HS, An J, et al. Ten-Eleven-Translocation 2 (TET2) negatively regulates homeostasis and differentiation of hematopoietic stem cells in mice. *Proc Natl Acad Sci*. 2011;108(35):14566-14571. doi:10.1073/pnas.1112317108.

233. Ko M, An J, Pastor WA, Koralov SB, Rajewsky K, Rao A. TET proteins and 5-methylcytosine oxidation in hematological cancers. *Immunol Rev*. 2015;263(1):6-21. doi:10.1111/imr.12239.
234. Figueroa ME, Abdel-Wahab O, Lu C, et al. Leukemic IDH1 and IDH2 Mutations Result in a Hypermethylation Phenotype, Disrupt TET2 Function, and Impair Hematopoietic Differentiation. *Cancer Cell*. 2010;18(6):553-567. doi:10.1016/j.ccr.2010.11.015.
235. Dang L, White DW, Gross S, et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature*. 2009;462(7274):739-744. doi:10.1038/nature08617.
236. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet*. 2017;18(9):517-534. doi:10.1038/nrg.2017.33.
237. Johnson KC, Houseman EA, King JE, von Herrmann KM, Fadul CE, Christensen BC. 5-Hydroxymethylcytosine localizes to enhancer elements and is associated with survival in glioblastoma patients. *Nat Commun*. 2016;7:13177. doi:10.1038/ncomms13177.
238. Lian CG, Xu Y, Ceol C, et al. Loss of 5-hydroxymethylcytosine is an epigenetic hallmark of melanoma. *Cell*. 2012;150(6):1135-1146. doi:10.1016/j.cell.2012.07.033.
239. Kudo Y, Tateishi K, Yamamoto K, et al. Loss of 5-hydroxymethylcytosine is accompanied with malignant cellular transformation. *Cancer Sci*. 2012;103(4):670-676. doi:10.1111/j.1349-7006.2012.02213.x.
240. Szabo PE, Pfeifer GP. H3K9me2 attracts PGC7 in the zygote to prevent Tet3-mediated oxidation of 5-methylcytosine. *J Mol Cell Biol*. 2012;4(6):427-429. doi:10.1093/jmcb/mjs038.
241. Iqbal K, Jin S-G, Pfeifer GP, Szabó PE. Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine. *Proc Natl Acad Sci*. 2011;108(9):3642-3647. doi:10.1073/pnas.1014033108.
242. Gu T-P, Guo F, Yang H, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 2011;477(7366):606-610. doi:10.1038/nature10443.
243. Wossidlo M, Nakamura T, Lepikhov K, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. *Nat Commun*. 2011;2:241. doi:10.1038/ncomms1240.
244. Inoue A, Shen L, Matoba S, Zhang Y. Haploinsufficiency, but Not Defective Paternal 5mC Oxidation, Accounts for the Developmental Defects of Maternal Tet3 Knockouts. *Cell Rep*. 2015;10(4):463-470. doi:10.1016/j.celrep.2014.12.049.
245. Yamaguchi S, Hong K, Liu R, et al. Dynamics of 5-methylcytosine and 5-hydroxymethylcytosine during germ cell reprogramming. *Cell Res*. 2013;23(3):329-339. doi:10.1038/cr.2013.22.
246. Kagiwada S, Kurimoto K, Hirota T, Yamaji M, Saitou M. Replication-coupled passive DNA demethylation for the erasure of genome imprints in mice. *EMBO J*. 2013;32(3):340-353. doi:10.1038/emboj.2012.331.
247. Yamaguchi S, Shen L, Liu Y, Sendler D, Zhang Y. Role of Tet1 in erasure of genomic imprinting. *Nature*. 2013;504(7480):460-464. doi:10.1038/nature12805.
248. Dawlaty MM, Breiling A, Le T, et al. Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. *Dev Cell*. 2013;24(3):310-323. doi:10.1016/j.devcel.2012.12.015.
249. Ruzov A, Tsenkina Y, Serio A, et al. Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. *Cell Res*. 2011;21(9):1332-1342. doi:10.1038/cr.2011.113.
250. Hentze H, Soong PL, Wang ST, Phillips BW, Putti TC, Dunn NR. Teratoma formation by human embryonic stem cells: Evaluation of essential parameters for future safety studies. *Stem Cell Res*. 2009;2(3):198-210. doi:10.1016/j.scr.2009.02.002.
251. Dawlaty MM, Ganz K, Powell BE, et al. Tet1 Is Dispensable for Maintaining Pluripotency and Its Loss Is Compatible with Embryonic and Postnatal

Development. *Cell Stem Cell*. 2011;9(2):166-175. doi:10.1016/j.stem.2011.07.010.
252. Koh KP, Yabuuchi A, Rao S, et al. Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell*. 2011;8(2):200-213. doi:10.1016/j.stem.2011.01.008.
253. Li Z, Cai X, Cai C-L, et al. Deletion of Tet2 in mice leads to dysregulated hematopoietic stem cells and subsequent development of myeloid malignancies. *Blood*. 2011;118(17):4509-4518. doi:10.1182/blood-2010-12-325241.
254. Dawlaty MM, Breiling A, Le T, et al. Loss of Tet enzymes compromises proper differentiation of embryonic stem cells. *Dev Cell*. 2014;29(1):102-111. doi:10.1016/j.devcel.2014.03.003.
255. Hu X, Zhang L, Mao S-Q, et al. Tet and TDG Mediate DNA Demethylation Essential for Mesenchymal-to-Epithelial Transition in Somatic Cell Reprogramming. *Cell Stem Cell*. 2014;14(4):512-522. doi:10.1016/j.stem.2014.01.001.
256. Lu F, Liu Y, Jiang L, Yamaguchi S, Zhang Y. Role of Tet proteins in enhancer activity and telomere elongation. *Genes Dev*. 2014;28(19):2103-2119. doi:10.1101/gad.248005.114.
257. Cortázar D, Kunz C, Selfridge J, et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature*. 2011;470(7334):419-423. doi:10.1038/nature09672.
258. Dai H-Q, Wang B-A, Yang L, et al. TET-mediated DNA demethylation controls gastrulation by regulating Lefty–Nodal signalling. *Nature*. 2016;538(7626):528-532. doi:10.1038/nature20095.
259. Kriaucionis S, Heintz N. The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*. 2009;324(5929):929-930. doi:10.1126/science.1169786.
260. Globisch D, Münzel M, Müller M, et al. Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. Croft AK, ed. *PLoS One*. 2010;5(12):e15367. doi:10.1371/journal.pone.0015367.
261. Santiago M, Antunes C, Guedes M, Sousa N, Marques CJ. TET enzymes and DNA hydroxymethylation in neural development and function — How critical are they? *Genomics*. 2014;104(5):334-340. doi:10.1016/j.ygeno.2014.08.018.
262. Nanduri J, Semenza GL, Prabhakar NR. Epigenetic changes by DNA methylation in chronic and intermittent hypoxia. *Am J Physiol Cell Mol Physiol*. 2017;313(6):L1096-L1100. doi:10.1152/ajplung.00325.2017.
263. Prabhakar NR, Semenza GL. Adaptive and Maladaptive Cardiorespiratory Responses to Continuous and Intermittent Hypoxia Mediated by Hypoxia-Inducible Factors 1 and 2. *Physiol Rev*. 2012;92(3):967-1003. doi:10.1152/physrev.00030.2011.
264. Chen H, Yan Y, Davidson TL, Shinkai Y, Costa M. Hypoxic Stress Induces Dimethylated Histone H3 Lysine 9 through Histone Methyltransferase G9a in Mammalian Cells. *Cancer Res*. 2006;66(18):9009-9016. doi:10.1158/0008-5472.CAN-06-0101.
265. Tachibana M, Sugimoto K, Nozaki M, et al. G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev*. 2002;16(14):1779-1791. doi:10.1101/gad.989402.
266. Cloos PAC, Christensen J, Agger K, Helin K. Erasing the methyl mark: histone demethylases at the center of cellular differentiation and disease. *Genes Dev*. 2008;22(9):1115-1140. doi:10.1101/gad.1652908.
267. Hancock RL, Dunne K, Walport LJ, Flashman E, Kawamura A. Epigenetic regulation by histone demethylases in hypoxia. *Epigenomics*. 2015;7(5):791-811. doi:10.2217/epi.15.24.
268. Johnson AB, Denko N, Barton MC. Hypoxia induces a novel signature of chromatin modifications and global repression of transcription. *Mutat Res*.

2008;640(1-2):174-179. doi:10.1016/j.mrfmmm.2008.01.001.
269. Pollard PJ, Loenarz C, Mole DR, et al. Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1α. *Biochem J*. 2008;416(3):387-394. doi:10.1042/BJ20081238.
270. Lee H-Y, Choi K, Oh H, Park Y-K, Park H. HIF-1-dependent induction of Jumonji domain-containing protein (JMJD) 3 under hypoxic conditions. *Mol Cells*. 2014;37(1):43-50. doi:10.14348/molcells.2014.2250.
271. Xia X, Lemieux ME, Li W, et al. Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. *Proc Natl Acad Sci*. 2009;106(11):4260-4265. doi:10.1073/pnas.0810067106.
272. Melvin A, Rocha S. Chromatin as an oxygen sensor and active player in the hypoxia response. *Cell Signal*. 2012;24(1):35-43. doi:10.1016/j.cellsig.2011.08.019.
273. Cascella B, Mirica LM. Kinetic Analysis of Iron-Dependent Histone Demethylases: α-Ketoglutarate Substrate Inhibition and Potential Relevance to the Regulation of Histone Demethylation in Cancer Cells. *Biochemistry*. 2012;51(44):8699-8701. doi:10.1021/bi3012466.
274. Hancock RL, Masson N, Dunne K, Flashman E, Kawamura A. The Activity of JmjC Histone Lysine Demethylase KDM4A is Highly Sensitive to Oxygen Concentrations. *ACS Chem Biol*. 2017;12(4):1011-1019. doi:10.1021/acschembio.6b00958.
275. Kasper LH, Boussouar F, Boyd K, et al. Two transactivation mechanisms cooperate for the bulk of HIF-1-responsive gene expression. *EMBO J*. 2005;24(22):3846-3858. doi:10.1038/sj.emboj.7600846.
276. Wang F, Zhang R, Wu X, Hankinson O. Roles of Coactivators in Hypoxic Induction of the Erythropoietin Gene. Pazin MJ, ed. *PLoS One*. 2010;5(4):e10002. doi:10.1371/journal.pone.0010002.
277. Jung JE, Lee HG, Cho IH, et al. STAT3 is a potential modulator of HIF-1-mediated VEGF expression in human renal carcinoma cells. *FASEB J*. 2005;19(10):1296-1298. doi:10.1096/fj.04-3099fje.
278. Kim MS, Kwon HJ, Lee YM, et al. Histone deacetylases induce angiogenesis by negative regulation of tumor suppressor genes. *Nat Med*. 2001;7(4):437-443. doi:10.1038/86507.
279. Kwon HJ, Kim MS, Kim MJ, Nakajima H, Kim K-W. Histone deacetylase inhibitor FK228 inhibits tumor angiogenesis. *Int J cancer*. 2002;97(3):290-296.
280. Kato H, Tamamizu-Kato S, Shibasaki F. Histone Deacetylase 7 Associates with Hypoxia-inducible Factor 1 and Increases Transcriptional Activity. *J Biol Chem*. 2004;279(40):41966-41974. doi:10.1074/jbc.M406320200.
281. Wu M-Z, Tsai Y-P, Yang M-H, et al. Interplay between HDAC3 and WDR5 Is Essential for Hypoxia-Induced Epithelial-Mesenchymal Transition. *Mol Cell*. 2011;43(5):811-822. doi:10.1016/j.molcel.2011.07.012.
282. Tsai Y-P, Chen H-F, Chen S-Y, et al. TET1 regulates hypoxia-induced epithelial-mesenchymal transition by acting as a co-activator. *Genome Biol*. 2014;15(12):513. doi:10.1186/s13059-014-0513-0.
283. Chen H-F, Wu K-J. Epigenetics, TET proteins, and hypoxia in epithelial-mesenchymal transition and tumorigenesis. *BioMedicine*. 2016;6(1):1. doi:10.7603/s40681-016-0001-9.
284. Shahrzad S, Bertrand K, Minhas K, Coomber BL. Induction of DNA hypomethylation by tumor hypoxia. *Epigenetics*. 2007;2(2):119-125.
285. Liu Q, Liu L, Zhao Y, et al. Hypoxia Induces Genomic DNA Demethylation through the Activation of HIF-1 and Transcriptional Upregulation of MAT2A in Hepatoma Cells. *Mol Cancer Ther*. 2011;10(6):1113-1123. doi:10.1158/1535-7163.MCT-10-1010.
286. Koslowski M, Luxemburger U, Türeci Ö, Sahin U. Tumor-associated CpG demethylation augments hypoxia-induced effects by positive autoregulation of HIF-1α. *Oncogene*. 2011;30(7):876-882. doi:10.1038/onc.2010.481.
287. Watson JA, Watson CJ, McCrohan A-M, et al. Generation of an epigenetic

signature by chronic hypoxia in prostate cells. *Hum Mol Genet*. 2009;18(19):3594-3604. doi:10.1093/hmg/ddp307.
288. Watson CJ, Collier P, Tea I, et al. Hypoxia-induced epigenetic modifications are associated with cardiac tissue fibrosis and the development of a myofibroblast-like phenotype. *Hum Mol Genet*. 2014;23(8):2176-2188. doi:10.1093/hmg/ddt614.
289. Skowronski K, Dubey S, Rodenhiser D, Coomber B. Ischemia dysregulates DNA methyltransferases and p16INK4a methylation in human colorectal cancer cells. *Epigenetics*. 2010;5(6):547-556.
290. Han X, Zhou Y, You Y, et al. TET1 promotes cisplatin-resistance via demethylating the vimentin promoter in ovarian cancer. *Cell Biol Int*. 2017;41(4):405-414. doi:10.1002/cbin.10734.
291. Tian Y, Pan F, Sun X, et al. Association of TET1 expression with colorectal cancer progression. *Scand J Gastroenterol*. 2017;52(3):312-320. doi:10.1080/00365521.2016.1253767.
292. Kienhöfer S, Musheev MU, Stapf U, et al. GADD45a physically and functionally interacts with TET1. *Differentiation*. 2015;90(1-3):59-68. doi:10.1016/j.diff.2015.10.003.
293. Zhang W, Lu Z, Gao Y, Ye L, Song T, Zhang X. MiR-520b suppresses proliferation of hepatoma cells through targeting ten-eleven translocation 1 (TET1) mRNA. *Biochem Biophys Res Commun*. 2015;460(3):793-798. doi:10.1016/j.bbrc.2015.03.108.
294. Moore CB, Guthrie EH, Tze- M, Huang H, Taxman DJ. Short Hairpin RNA (shRNA): Design, Delivery, and Assessment of Gene Knockdown. doi:10.1007/978-1-60761-657-3_10.
295. Brown T. Dot and slot blotting of DNA. *Curr Protoc Mol Biol*. 2001;Chapter 2:Unit2.9B. doi:10.1002/0471142727.mb0209bs21.
296. Le T, Kim K-P, Fan G, Faull KF. A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. *Anal Biochem*. 2011;412(2):203-209. doi:10.1016/j.ab.2011.01.026.
297. Ludwig C, Günther UL. MetaboLab - advanced NMR data processing and analysis for metabolomics. *BMC Bioinformatics*. 2011;12(1):366. doi:10.1186/1471-2105-12-366.
298. Larue L, Antos C, Butz S, et al. A role for cadherins in tissue formation. *Development*. 1996;122(10):3185-3194.
299. Dang SM, Gerecht-Nir S, Chen J, Itskovitz-Eldor J, Zandstra PW. Controlled, Scalable Embryonic Stem Cell Differentiation Culture. *Stem Cells*. 2004;22(3):275-282. doi:10.1634/stemcells.22-3-275.
300. Bratt-Leal AM, Carpenedo RL, McDevitt TC. Engineering the embryoid body microenvironment to direct embryonic stem cell differentiation. *Biotechnol Prog*. 2009;25(1):43-51. doi:10.1002/btpr.139.
301. Turing AM. The Chemical Basis of Morhpogenesis. *Philos Trans R Soc B Biol Sci*. 1952;237(641):37-72.
302. Ploumakis A, Coleman ML. OH, the Places You'll Go! Hydroxylation, Gene Expression, and Cancer. *Mol Cell*. 2015;58(5):729-741. doi:10.1016/j.molcel.2015.05.026.
303. Laukka T, Mariani CJ, Ihantola T, et al. Fumarate and Succinate Regulate Expression of Hypoxia-inducible Genes via TET Enzymes. *J Biol Chem*. 2016;291(8):4256-4265. doi:10.1074/jbc.M115.688762.
304. Hirsilä M, Koivunen P, Günzler V, Kivirikko KI, Myllyharju J. Characterization of the Human Prolyl 4-Hydroxylases That Modify the Hypoxia-inducible Factor. *J Biol Chem*. 2003;278(33):30772-30780. doi:10.1074/jbc.M304982200.
305. Reynafarje B, Costa LE, Lehninger AL. O2 solubility in aqueous media determined by a kinetic method. *Anal Biochem*. 1985;145(2):406-418.
306. Schofield CJ, Ratcliffe PJ. Signalling hypoxia by HIF hydroxylases. 2005. doi:10.1016/j.bbrc.2005.08.111.

307. Fandrey J, Gorr T, Gassmann M. Regulating cellular oxygen sensing by hydroxylation. *Cardiovasc Res*. 2006;71(4):642-651. doi:10.1016/j.cardiores.2006.05.005.
308. Elkins JM, Hewitson KS, McNeill LA, et al. Structure of Factor-inhibiting Hypoxia-inducible Factor (HIF) Reveals Mechanism of Oxidative Modification of HIF-1α. *J Biol Chem*. 2003;278(3):1802-1806. doi:10.1074/jbc.C200644200.
309. Thienpont B, Steinbacher J, Zhao H, et al. Tumour hypoxia causes DNA hypermethylation by reducing TET activity Effect of hypoxia on DNA hydroxymethylation. 2016. doi:10.1038/nature19081.
310. Astuti D, Latif F, Dallol A, et al. Gene mutations in the succinate dehydrogenase subunit SDHB cause susceptibility to familial pheochromocytoma and to familial paraganglioma. *Am J Hum Genet*. 2001;69(1):49-54. doi:10.1086/321282.
311. Tomlinson IPM, Alam NA, Rowan AJ, et al. Germline mutations in FH predispose to dominantly inherited uterine fibroids, skin leiomyomata and papillary renal cell cancer. *Nat Genet*. 2002;30(4):406-410. doi:10.1038/ng849.
312. Xiao M, Yang H, Xu W, et al. Inhibition of a-KG-dependent histone and DNA demethylases by fumarate and succinate that are accumulated in mutations of FH and SDH tumor suppressors. doi:10.1101/gad.191056.112.
313. Yang M, Pollard PJ, Adam J, et al. Succinate: A New Epigenetic Hacker. *Cancer Cell*. 2013;23(6):709-711. doi:10.1016/j.ccr.2013.05.015.
314. Selak MA, Armour SM, MacKenzie ED, et al. Succinate links TCA cycle dysfunction to oncogenesis by inhibiting HIF-α prolyl hydroxylase. *Cancer Cell*. 2005;7(1):77-85. doi:10.1016/j.ccr.2004.11.022.
315. Isaacs JS, Jung YJ, Mole DR, et al. HIF overexpression correlates with biallelic loss of fumarate hydratase in renal cancer: Novel role of fumarate in regulation of HIF stability. *Cancer Cell*. 2005;8(2):143-153. doi:10.1016/j.ccr.2005.06.017.
316. King A, Selak MA, Gottlieb E. Succinate dehydrogenase and fumarate hydratase: linking mitochondrial dysfunction and cancer. *Oncogene*. 2006;25(34):4675-4682. doi:10.1038/sj.onc.1209594.
317. Aragonés J, Fraisl P, Baes M, Carmeliet P. Oxygen Sensors at the Crossroad of Metabolism. *Cell Metab*. 2009;9(1):11-22. doi:10.1016/j.cmet.2008.10.001.
318. Powers DE, Millman JR, Huang RB, Colton CK. Effects of oxygen on mouse embryonic stem cell growth, phenotype retention, and cellular energetics. *Biotechnol Bioeng*. 2008;101(2):241-254. doi:10.1002/bit.21986.
319. Benzi G, Arrigoni E, Marzatico F, Villa RF. Influence of some biological pyrimidines on the succinate cycle during and after cerebral ischemia. *Biochem Pharmacol*. 1979;28(17):2545-2550.
320. Folbergrová J, Ljunggren B, Norberg K, Siesjö BK. Influence of complete ischemia on glycolytic metabolites, citric acid cycle intermediates, and associated amino acids in the rat cerebral cortex. *Brain Res*. 1974;80(2):265-279.
321. Li T, Yang D, Li J, Tang Y, Yang J, Le W. Critical Role of Tet3 in Neural Progenitor Cell Maintenance and Terminal Differentiation. *Mol Neurobiol*. 2015;51(1):142-154. doi:10.1007/s12035-014-8734-5.
322. Li X, Yue X, Pastor WA, et al. Tet proteins influence the balance between neuroectodermal and mesodermal fate choice by inhibiting Wnt signaling. *Proc Natl Acad Sci U S A*. 2016;113(51):E8267-E8276. doi:10.1073/pnas.1617802113.
323. Huang Y, Chavez L, Chang X, et al. Distinct roles of the methylcytosine oxidases Tet1 and Tet2 in mouse embryonic stem cells. *Proc Natl Acad Sci U S A*. 2014;111(4):1361-1366. doi:10.1073/pnas.1322921111.
324. Langlois T, da Costa Reis Monte-Mor B, Lenglet G, et al. TET2 Deficiency Inhibits Mesoderm and Hematopoietic Differentiation in Human Embryonic Stem Cells. *Stem Cells*. 2014;32(8):2084-2097. doi:10.1002/stem.1718.
325. Wang Z, Tang B, He Y, Jin P. DNA methylation dynamics in neurogenesis. *Epigenomics*. 2016;8(3):401-414. doi:10.2217/epi.15.119.
326. Yan H, Wang Y, Qu X, et al. Distinct roles for TET family proteins in regulating

human erythropoiesis. *Blood*. 2017;129(14):2002-2012. doi:10.1182/blood-2016-08-736587.
327. Xu Y, Xu C, Kato A, et al. Tet3 CXXC domain and dioxygenase activity cooperatively regulate key genes for Xenopus eye and neural development. *Cell*. 2012;151(6):1200-1213. doi:10.1016/j.cell.2012.11.014.
328. van der Heyden MAG, Defize LHK. Twenty one years of P19 cells: what an embryonal carcinoma cell line taught us about cardiomyocyte differentiation. *Cardiovasc Res*. 2003;58(2):292-302.
329. Mcburney M. P19 embryonal carcinoma cells. 1993;140:135-140.
330. Pesce M, Schöler HR. *Oct-4* : Gatekeeper in the Beginnings of Mammalian Development. *Stem Cells*. 2001;19(4):271-278. doi:10.1634/stemcells.19-4-271.
331. Pesce M, Scholer HR. Oct-4: Control of totipotency and germline determination. *Mol Reprod Dev*. 2000;55(4):452-457. doi:10.1002/(SICI)1098-2795(200004)55:4<452::AID-MRD14>3.0.CO;2-S.
332. Chambers I, Colby D, Robertson M, et al. Functional expression cloning of Nanog, a pluripotency sustaining factor in embryonic stem cells. *Cell*. 2003;113(5):643-655.
333. Mitsui K, Tokuzawa Y, Itoh H, et al. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell*. 2003;113(5):631-642.
334. Silva J, Nichols J, Theunissen TW, et al. Nanog is the gateway to the pluripotent ground state. *Cell*. 2009;138(4):722-737. doi:10.1016/j.cell.2009.07.039.
335. Burtscher I, Lickert H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development*. 2009;136(6):1029-1038. doi:10.1242/dev.028415.
336. Futaki S, Hayashi Y, Emoto T, Weber CN, Sekiguchi K. Sox7 plays crucial roles in parietal endoderm differentiation in F9 embryonal carcinoma cells through regulating Gata-4 and Gata-6 expression. *Mol Cell Biol*. 2004;24(23):10492-10503. doi:10.1128/MCB.24.23.10492-10503.2004.
337. Kim PTW, Hoffman BG, Plesner A, et al. Differentiation of Mouse Embryonic Stem Cells into Endoderm without Embryoid Body Formation. Maedler K, ed. *PLoS One*. 2010;5(11):e14146. doi:10.1371/journal.pone.0014146.
338. Herrmann BG, Labeit S, Poustka A, King TR, Lehrach H. Cloning of the T gene required in mesoderm formation in the mouse. *Nature*. 1990;343(6259):617-622. doi:10.1038/343617a0.
339. Riley P, Anaon-Cartwight L, Cross JC. The Hand1 bHLH transcription factor is essential for placentation and cardiac morphogenesis. *Nat Genet*. 1998;18(3):271-275. doi:10.1038/ng0398-271.
340. Chan SS-K, Shi X, Toyama A, et al. Mesp1 patterns mesoderm into cardiac, hematopoietic, or skeletal myogenic progenitors in a context-dependent manner. *Cell Stem Cell*. 2013;12(5):587-601. doi:10.1016/j.stem.2013.03.004.
341. Horb ME, Thomsen GH. Tbx5 is essential for heart development. *Development*. 1999;126(8):1739-1751.
342. Watt AJ, Battle MA, Li J, Duncan SA. GATA4 is essential for formation of the proepicardium and regulates cardiogenesis. *Proc Natl Acad Sci U S A*. 2004;101(34):12573-12578. doi:10.1073/pnas.0400752101.
343. Dráberová E, Del Valle L, Gordon J, et al. Class III β-Tubulin Is Constitutively Coexpressed With Glial Fibrillary Acidic Protein and Nestin in Midgestational Human Fetal Astrocytes: Implications for Phenotypic Identity. *J Neuropathol Exp Neurol*. 2008;67(4):341-354. doi:10.1097/NEN.0b013e31816a686d.
344. Rønn LC, Hartz BP, Bock E. The neural cell adhesion molecule (NCAM) in development and plasticity of the nervous system. *Exp Gerontol*. 33(7-8):853-864.
345. Hendrickson ML, Rao AJ, Demerdash ONA, Kalil RE. Expression of Nestin by Neural Cells in the Adult Rat and Human Brain. Combs C, ed. *PLoS One*. 2011;6(4):e18535. doi:10.1371/journal.pone.0018535.
346. Landgren H, Carlsson P. Foxj3, a novel mammalian forkhead gene expressed in

neuroectoderm, neural crest, and myotome. *Dev Dyn*. 2004;231(2):396-401. doi:10.1002/dvdy.20131.
347. Iwata M, Muneoka KT, Shirayama Y, Yamamoto A, Kawahara R. A study of a dendritic marker, microtubule-associated protein 2 (MAP-2), in rats neonatally treated neurosteroids, pregnenolone and dehydroepiandrosterone (DHEA). *Neurosci Lett*. 2005;386(3):145-149. doi:10.1016/j.neulet.2005.06.004.
348. Asikainen TM, Ahmad A, Schneider BK, et al. Stimulation of HIF-1α, HIF-2α, and VEGF by prolyl 4-hydroxylase inhibition in human lung endothelial and epithelial cells. *Free Radic Biol Med*. 2005;38(8):1002-1013. doi:10.1016/j.freeradbiomed.2004.12.004.
349. Ying Q-L, Wray J, Nichols J, et al. The ground state of embryonic stem cell self-renewal. *Nature*. 2008;453(7194):519-523. doi:10.1038/nature06968.
350. Chavez JC, Almhanna K, Berti-Mattera LN. Transient expression of hypoxia-inducible factor-1 alpha and target genes in peripheral nerves from diabetic rats. *Neurosci Lett*. 2005;374(3):179-182. doi:10.1016/j.neulet.2004.10.050.
351. Bagnall J, Leedale J, Taylor SE, et al. Tight Control of Hypoxia-inducible Factor-α Transient Dynamics Is Essential for Cell Survival in Hypoxia. *J Biol Chem*. 2014;289(9):5549-5564. doi:10.1074/jbc.M113.500405.
352. C. elegans EGL-9 and Mammalian Homologs Define a Family of Dioxygenases that Regulate HIF by Prolyl Hydroxylation. *Cell*. 2001;107(1):43-54. doi:10.1016/S0092-8674(01)00507-4.
353. Maeno M, Mead PE, Kelley C, et al. The role of BMP-4 and GATA-2 in the induction and differentiation of hematopoietic mesoderm in Xenopus laevis. *Blood*. 1996;88(6):1965-1972.
354. Gleeson JG, Lin PT, Flanagan LA, Walsh CA. Doublecortin is a microtubule-associated protein and is expressed widely by migrating neurons. *Neuron*. 1999;23(2):257-271.
355. Levine AJ, Brivanlou AH, Suzuki A, Hemmati-Brivanlou A. GDF3, a BMP inhibitor, regulates cell fate in stem cells and early embryos. *Development*. 2006;133(2):209-216. doi:10.1242/dev.02192.
356. Houweling A, Vanborren M, Morrman A, Christoffels V. Expression and regulation of the atrial natriuretic factor encoding gene during development and disease. *Cardiovasc Res*. 2005;67(4):583-593. doi:10.1016/j.cardiores.2005.06.013.
357. Mahdavi V, Periasamy M, Nadal-Ginard B. Molecular characterization of two myosin heavy chain genes expressed in the adult heart. *Nature*. 1982;297(5868):659-664.
358. Boutin C, Hardt O, de Chevigny A, et al. NeuroD1 induces terminal neuronal differentiation in olfactory neurogenesis. *Proc Natl Acad Sci U S A*. 2010;107(3):1201-1206. doi:10.1073/pnas.0909015107.
359. Kim D-K, Cha Y, Ahn H-J, Kim G, Park K-S. *Lefty1* and *Lefty2* Control the Balance Between Self-Renewal and Pluripotent Differentiation of Mouse Embryonic Stem Cells. *Stem Cells Dev*. 2014;23(5):457-466. doi:10.1089/scd.2013.0220.
360. Pawlak M, Jaenisch R. De novo DNA methylation by Dnmt3a and Dnmt3b is dispensable for nuclear reprogramming of somatic cells to a pluripotent state. *Genes Dev*. 2011;25(10):1035-1040. doi:10.1101/gad.2039011.
361. Wang P, McKnight KD, Wong DJ, et al. A Molecular Signature for Purified Definitive Endoderm Guides Differentiation and Isolation of Endoderm from Mouse and Human Embryonic Stem Cells. *Stem Cells Dev*. 2012;21(12):2273. doi:10.1089/SCD.2011.0416.
362. Zhang C, Ye X, Zhang H, Ding M, Deng H. GATA Factors Induce Mouse Embryonic Stem Cell Differentiation Toward Extraembryonic Endoderm. *Stem Cells Dev*. 2007;16(4):605-614. doi:10.1089/scd.2006.0077.
363. Dupuis LE, Kern CB. Small leucine-rich proteoglycans exhibit unique spatiotemporal expression profiles during cardiac valve development. *Dev Dyn*. 2014;243(4):601-611. doi:10.1002/dvdy.24100.

364. Elkabetz Y, Panagiotakos G, Al Shamy G, Socci ND, Tabar V, Studer L. Human ES cell-derived neural rosettes reveal a functionally distinct early neural stem cell stage. *Genes Dev*. 2008;22(2):152-165. doi:10.1101/gad.1616208.
365. Mallika C, Guo Q, Li JYH. Gbx2 is essential for maintaining thalamic neuron identity and repressing habenular characters in the developing thalamus. *Dev Biol*. 2015;407(1):26-39. doi:10.1016/j.ydbio.2015.08.010.
366. Cichorek M, Wachulska M, Stasiewicz A, Tymińska A. Skin melanocytes: biology and development. *Postep dermatologii i Alergol*. 2013;30(1):30-41. doi:10.5114/pdia.2013.33376.
367. Ohtsuka T, Sakamoto M, Guillemot F, Kageyama R. Roles of the Basic Helix-Loop-Helix Genes *Hes1* and *Hes5* in Expansion of Neural Stem Cells of the Developing Brain. *J Biol Chem*. 2001;276(32):30467-30474. doi:10.1074/jbc.M102420200.
368. Holmberg Olausson K, Maire CL, Haidar S, et al. Prominin-1 (CD133) Defines Both Stem and Non-Stem Cell Populations in CNS Development and Gliomas. Canoll P, ed. *PLoS One*. 2014;9(9):e106694. doi:10.1371/journal.pone.0106694.
369. Mizuguchi R, Sugimori M, Takebayashi H, et al. Combinatorial roles of olig2 and neurogenin2 in the coordinated induction of pan-neuronal and subtype-specific properties of motoneurons. *Neuron*. 2001;31(5):757-771.
370. Ma K, Fox L, Shi G, et al. Generation of neural stem cell-like cells from bone marrow-derived human mesenchymal stem cells. *Neurol Res*. 2011;33(10):1083-1093. doi:10.1179/1743132811Y.0000000053.
371. Eng LF. Glial fibrillary acidic protein (GFAP): the major protein of glial intermediate filaments in differentiated astrocytes. *J Neuroimmunol*. 1985;8(4-6):203-214.
372. Wu L-Y, Yue A, Ae W, et al. The Role of Hypoxia in the Differentiation of P19 Embryonal Carcinoma Cells into Dopaminergic Neurons. doi:10.1007/s11064-008-9728-3.
373. Lolmède K, Durand de Saint Front V, Galitzky J, Lafontan M, Bouloumié A. Effects of hypoxia on the expression of proangiogenic factors in differentiated 3T3-F442A adipocytes. *Int J Obes*. 2003;27(10):1187-1195. doi:10.1038/sj.ijo.0802407.
374. Gustafsson M V, Zheng X, Pereira T, et al. Hypoxia requires notch signaling to maintain the undifferentiated cell state. *Dev Cell*. 2005;9(5):617-628. doi:10.1016/j.devcel.2005.09.010.
375. Zhou X, Smith AJH, Waterhouse A, et al. Hes1 desynchronizes differentiation of pluripotent cells by modulating STAT3 activity. *Stem Cells*. 2013;31(8):1511-1522. doi:10.1002/stem.1426.
376. Medley TL, Furtado M, Lam NT, et al. Effect of oxygen on cardiac differentiation in mouse iPS cells: role of hypoxia inducible factor-1 and Wnt/beta-catenin signaling. Najbauer J, ed. *PLoS One*. 2013;8(11):e80280. doi:10.1371/journal.pone.0080280.
377. Mazumdar J, O 'brien WT, Johnson RS, et al. O2 regulates stem cells through Wnt/β-catenin signalling. 2010. doi:10.1038/ncb2102.
378. Sokol SY. Maintaining embryonic stem cell pluripotency with Wnt signaling. *Development*. 2011;138(20):4341-4350. doi:10.1242/dev.066209.
379. Couillard-Despres S, Winner B, Schaubeck S, et al. Doublecortin expression levels in adult brain reflect neurogenesis. *Eur J Neurosci*. 2005;21(1):1-14. doi:10.1111/j.1460-9568.2004.03813.x.
380. Studer L, Csete M, Lee SH, et al. Enhanced proliferation, survival, and dopaminergic differentiation of CNS precursors in lowered oxygen. *J Neurosci*. 2000;20(19):7377-7383.
381. Santilli G, Lamorte G, Carlessi L, et al. Mild Hypoxia Enhances Proliferation and Multipotency of Human Neural Stem Cells. Najbauer J, ed. *PLoS One*. 2010;5(1):e8575. doi:10.1371/journal.pone.0008575.
382. Kim T-S, Misumi S, Jung C-G, et al. Increase in dopaminergic neurons from mouse embryonic stem cell-derived neural progenitor/stem cells is mediated by

hypoxia inducible factor-1α. *J Neurosci Res*. 2008;86(11):2353-2362. doi:10.1002/jnr.21687.
383. Binh NH, Aoki H, Takamatsu M, et al. Time-sensitive effects of hypoxia on differentiation of neural stem cells derived from mouse embryonic stem cells *in vitro*. *Neurol Res*. 2014;36(9):804-813. doi:10.1179/1743132814Y.0000000338.
384. Chen T, Dent SYR. Chromatin modifiers and remodellers: regulators of cellular differentiation. *Nat Rev Genet*. 2013;15(2):93-106. doi:10.1038/nrg3607.
385. Tsai Y-P, Wu K-J. Epigenetic regulation of hypoxia-responsive gene expression: Focusing on chromatin and DNA modifications. *Int J Cancer*. 2014;134(2):249-256. doi:10.1002/ijc.28190.
386. Ito S, D'Alessio AC, Taranova O V., Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*. 2010;466(7310):1129-1133. doi:10.1038/nature09303.
387. Stroud H, Feng S, Morey Kinney S, Pradhan S, Jacobsen SE. 5-Hydroxymethylcytosine is associated with enhancers and gene bodies in human embryonic stem cells. *Genome Biol*. 2011;12(6):R54. doi:10.1186/gb-2011-12-6-r54.
388. Wu H, Zhang Y. Tet1 and 5-hydroxymethylation A genome-wide view in mouse embryonic stem cells. *Cell Cycle*. 2011;1015:2428-2436. doi:10.4161/cc.10.15.16930.
389. Williams K, Christensen J, Pedersen MT, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature*. 2011;473(7347):343-348. doi:10.1038/nature10066.
390. Wu H, D'Alessio AC, Ito S, et al. Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev*. 2011;25(7):679-684. doi:10.1101/gad.2036011.
391. Costa Y, Ding J, Theunissen TW, et al. NANOG-dependent function of TET1 and TET2 in establishment of pluripotency. *Nature*. 2013;495(7441):370-374. doi:10.1038/nature11925.
392. Gu T-P, Guo F, Yang H, et al. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature*. 2011;477(7366):606-610. doi:10.1038/nature10443.
393. Wu H, Zhang Y. Mechanisms and functions of Tet protein-mediated 5-methylcytosine oxidation. *Genes Dev*. 2011;25(23):2436-2452. doi:10.1101/gad.179184.111.
394. Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. *Nat Rev Mol Cell Biol*. 2005;6(5):376-385. doi:10.1038/nrm1644.
395. Suárez Y, Sessa WC. MicroRNAs as novel regulators of angiogenesis. *Circ Res*. 2009;104(4):442-454. doi:10.1161/CIRCRESAHA.108.191270.
396. Almeida MI, Nicoloso MS, Zeng L, et al. Strand-Specific miR-28-5p and miR-28-3p Have Distinct Effects in Colorectal Cancer Cells. *Gastroenterology*. 2012;142(4):886-896.e9. doi:10.1053/j.gastro.2011.12.047.
397. Cheng J, Guo S, Chen S, et al. An Extensive Network of TET2-Targeting MicroRNAs Regulates Malignant Hematopoiesis. *Cell Rep*. 2013;5(2):471-481. doi:10.1016/j.celrep.2013.08.050.
398. Dill H, Linder B, Fehr A, Fischer U. Intronic miR-26b controls neuronal differentiation by repressing its host transcript, ctdsp2. *Genes Dev*. 2012;26(1):25-30. doi:10.1101/gad.177774.111.
399. Fu X, Jin L, Wang X, et al. MicroRNA-26a targets ten eleven translocation enzymes and is regulated during pancreatic cell differentiation. *Proc Natl Acad Sci U S A*. 2013;110(44):17892-17897. doi:10.1073/pnas.1317397110.
400. Morita S, Horii T, Kimura M, Ochiya T, Tajima S, Hatada I. miR-29 Represses the Activities of DNA Methyltransferases and DNA Demethylases. *Int J Mol Sci*. 2013;14(7):14647-14658. doi:10.3390/ijms140714647.
401. Tu J, Ng SH, Luk ACS, et al. MicroRNA-29b/Tet1 regulatory axis epigenetically modulates mesendoderm differentiation in mouse embryonic stem cells. *Nucleic*

Acids Res. 2015;43(16):7805-7822. doi:10.1093/nar/gkv653.
402. Kanellopoulou C, Muljo SA, Kung AL, et al. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev*. 2005;19(4):489-501. doi:10.1101/gad.1248505.
403. Bernstein E, Kim SY, Carmell MA, et al. Dicer is essential for mouse development. *Nat Genet*. 2003;35(3):215-217. doi:10.1038/ng1253.
404. Lai X, Vera J. MicroRNA Clusters. In: *Encyclopedia of Systems Biology*. New York, NY: Springer New York; 2013:1310-1314. doi:10.1007/978-1-4419-9863-7_1121.
405. Lee YJ, Ramakrishna S, Chauhan H, Park WS, Hong S-H, Kim K-S. Dissecting microRNA-mediated regulation of stemness, reprogramming, and pluripotency. doi:10.1186/s13619-016-0028-0.
406. Greer Card DA, Hebbar PB, Li L, et al. Oct4/Sox2-Regulated miR-302 Targets Cyclin D1 in Human Embryonic Stem Cells. *Mol Cell Biol*. 2008;28(20):6426-6438. doi:10.1128/MCB.00359-08.
407. Marson A, Levine SS, Cole MF, et al. Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell*. 2008;134(3):521-533. doi:10.1016/j.cell.2008.07.020.
408. Hohjoh H, Fukushima T. Marked change in microRNA expression during neuronal differentiation of human teratocarcinoma NTera2D1 and mouse embryonal carcinoma P19 cells. *Biochem Biophys Res Commun*. 2007;362(2):360-367. doi:10.1016/j.bbrc.2007.07.189.
409. Chen L, Heikkinen L, Emily Knott K, Liang Y, Wong G. Evolutionary conservation and function of the human embryonic stem cell specific miR-302/367 cluster. *Comp Biochem Physiol Part D Genomics Proteomics*. 2015;16:83-98. doi:10.1016/j.cbd.2015.08.002.
410. Lin S-L, Chang DC, Lin C-H, Ying S-Y, Leu D, Wu DTS. Regulation of somatic cell reprogramming through inducible mir-302 expression. *Nucleic Acids Res*. 2011;39(3):1054-1065. doi:10.1093/nar/gkq850.
411. Hu S, Wilson KD, Ghosh Z, et al. MicroRNA-302 increases reprogramming efficiency via repression of NR2F2. *Stem Cells*. 2013;31(2):259-268. doi:10.1002/stem.1278.
412. Lee MR, Prasain N, Chae H-D, et al. Epigenetic regulation of NANOG by miR-302 cluster-MBD2 completes induced pluripotent stem cell reprogramming. *Stem Cells*. 2013;31(4):666-681. doi:10.1002/stem.1302.
413. Kuo C-H, Deng JH, Deng Q, Ying S-Y. A novel role of miR-302/367 in reprogramming. *Biochem Biophys Res Commun*. 2012;417(1):11-16. doi:10.1016/j.bbrc.2011.11.058.
414. Lin S-L. Mechanism of MicroRNA-Mediated Global DNA Demethylation in Human iPS Cells. In: *Advances in Stem Cell Research*. Totowa, NJ: Humana Press; 2012:117-134. doi:10.1007/978-1-61779-940-2_8.
415. Li H-L, Wei J-F, Fan L-Y, et al. miR-302 regulates pluripotency, teratoma formation and differentiation in stem cells via an AKT1/OCT4-dependent manner. *Cell Death Dis*. 2016;7(1):e2078. doi:10.1038/cddis.2015.383.
416. Fidalgo M, Huang X, Guallar D, et al. Zfp281 Coordinates Opposing Functions of Tet1 and Tet2 in Pluripotent States. *Cell Stem Cell*. 2016;19(3):355-369. doi:10.1016/j.stem.2016.05.025.
417. Parchem RJ, Moore N, Fish JL, et al. miR-302 Is Required for Timing of Neural Differentiation, Neural Tube Closure, and Embryonic Viability. *Cell Rep*. 2015;12(5):760-773. doi:10.1016/j.celrep.2015.06.074.
418. Yang S-L, Yang M, Herrlinger S, Liang C, Lai F, Chen J-F. MiR-302/367 regulate neural progenitor proliferation, differentiation timing, and survival in neurulation. *Dev Biol*. 2015;408(1):140-150. doi:10.1016/j.ydbio.2015.09.020.
419. Lv X, Jiang H, Liu Y, Lei X, Jiao J. MicroRNA-15b promotes neurogenesis and inhibits neural progenitor proliferation by directly repressing TET3 during early neocortical development. *EMBO Rep*. 2014;15(12):1305-1314. doi:10.15252/embr.201438923.

420. Han L, Witmer PD, Casey E, Valle D, Sukumar S. DNA methylation regulates MicroRNA expression. *Cancer Biol Ther*. 2007;6(8):1284-1288.
421. Gulyaeva LF, Kushlinskiy NE. Regulatory mechanisms of microRNA expression. *J Transl Med*. 2016;14(1):143. doi:10.1186/s12967-016-0893-x.
422. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res*. 2010;38(19):e181. doi:10.1093/nar/gkq684.
423. Kinney SM, Chin HG, Vaisvila R, et al. Tissue-specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes. *J Biol Chem*. 2011;286(28):24685-24693. doi:10.1074/jbc.M110.217083.
424. Lind J. *A Treatise on the Scurvy : In Three Parts, Containing an Inquiry into the Nature, Causes, and Cure, of That Disease : Together with a Critical and Chronological View of What Has Been Published on the Subject*. London: Millar; 1757.
425. Svirbely JL, Szent-Györgyi A. The chemical nature of vitamin C. *Biochem J*. 1933;27(1):279-285.
426. Whitehead P. The history of scurvy and vitamin C. *Med Hist*. 1987;31(2):231.
427. Gorres KL, Raines RT. Prolyl 4-hydroxylase. *Crit Rev Biochem Mol Biol*. 2010;45(2):106-124. doi:10.3109/10409231003627991.
428. Myllylä R, Kuutti-Savolainen E-R, Kivirikko KI. The role of ascorbate in the prolyl hydroxylase reaction. *Biochem Biophys Res Commun*. 1978;83(2):441-448. doi:10.1016/0006-291X(78)91010-0.
429. Kuiper C, Dachs GU, Currie MJ, Vissers MCM. Intracellular ascorbate enhances hypoxia-inducible factor (HIF)-hydroxylase activity and preferentially suppresses the HIF-1 transcriptional response. *Free Radic Biol Med*. 2014;69:308-317. doi:10.1016/j.freeradbiomed.2014.01.033.
430. Nytko KJ, Maeda N, Schläfli P, Spielmann P, Wenger RH, Stiehl DP. Vitamin C is dispensable for oxygen sensing in vivo. *Blood*. 2011;117(20):5485-5493. doi:10.1182/blood-2010-09-307637.
431. Stadtfeld M, Apostolou E, Ferrari F, et al. Ascorbic acid prevents loss of Dlk1-Dio3 imprinting and facilitates generation of all–iPS cell mice from terminally differentiated B cells. *Nat Genet*. 2012;44(4):398-405. doi:10.1038/ng.1110.
432. Chung T-L, Brena RM, Kolle G, et al. Vitamin C Promotes Widespread Yet Specific DNA Demethylation of the Epigenome in Human Embryonic Stem Cells. *Stem Cells*. 2010;28(10):1848-1855. doi:10.1002/stem.493.
433. Chung T-L, Turner JP, Thaker NY, et al. Ascorbate Promotes Epigenetic Activation of CD30 in Human Embryonic Stem Cells. *Stem Cells*. 2010;28(10):1782-1793. doi:10.1002/stem.500.
434. Yin R, Mao S-Q, Zhao B, et al. Ascorbic Acid Enhances Tet-Mediated 5-Methylcytosine Oxidation and Promotes DNA Demethylation in Mammals. *J Am Chem Soc*. 2013;135(28):10396-10403. doi:10.1021/ja4028346.
435. Blaschke K, Ebata KT, Karimi MM, et al. Vitamin C induces Tet-dependent DNA demethylation and a blastocyst-like state in ES cells. *Nature*. 2013;500(7461):222-226. doi:10.1038/nature12362.
436. Chen J, Guo L, Zhang L, et al. Vitamin C modulates TET1 function during somatic cell reprogramming. *Nat Genet*. 2013;45(12):1504-1509. doi:10.1038/ng.2807.
437. Hore TA, von Meyenn F, Ravichandran M, et al. Retinol and ascorbate drive erasure of epigenetic memory and enhance reprogramming to naïve pluripotency by complementary mechanisms. *Proc Natl Acad Sci U S A*. 2016;113(43):12202-12207. doi:10.1073/pnas.1608679113.
438. Yan J, Studer L, McKay RD. Ascorbic acid increases the yield of dopaminergic neurons derived from basic fibroblast growth factor expanded mesencephalic precursors. *J Neurochem*. 2001;76(1):307-311.
439. He X-B, Kim M, Kim S-Y, et al. Vitamin C Facilitates Dopamine Neuron Differentiation in Fetal Midbrain Through TET1- and JMJD3-Dependent Epigenetic Control Manner. *Stem Cells*. 2015;33(4):1320-1332.

doi:10.1002/stem.1932.
440. Yue X, Trifari S, Äijö T, et al. Control of Foxp3 stability through modulation of TET activity. *J Exp Med*. 2016;213(3):377-397. doi:10.1084/jem.20151438.
441. Minor EA, Court BL, Young JI, Wang G. Ascorbate induces ten-eleven translocation (Tet) methylcytosine dioxygenase-mediated generation of 5-hydroxymethylcytosine. *J Biol Chem*. 2013;288(19):13669-13674. doi:10.1074/jbc.C113.464800.
442. Zhang S, Roche K, Nasheuer H-P, Lowndes NF. Modification of Histones by Sugar β- *N* -Acetylglucosamine (GlcNAc) Occurs on Multiple Residues, Including Histone H3 Serine 10, and Is Cell Cycle-regulated. *J Biol Chem*. 2011;286(43):37483-37495. doi:10.1074/jbc.M111.284885.
443. Fong JJ, Nguyen BL, Bridger R, et al. β-N-Acetylglucosamine (O-GlcNAc) Is a Novel Regulator of Mitosis-specific Phosphorylations on Histone H3. *J Biol Chem*. 2012;287(15):12195-12203. doi:10.1074/jbc.M111.315804.
444. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem*. 2012;81:65-95. doi:10.1146/annurev-biochem-051710-134100.
445. Aranda S, Mas G, Di Croce L. Regulation of gene transcription by Polycomb proteins. *Sci Adv*. 2015;1(11):e1500737-e1500737. doi:10.1126/sciadv.1500737.
446. Margueron R, Reinberg D. The Polycomb complex PRC2 and its mark in life. *Nature*. 2011;469(7330):343-349. doi:10.1038/nature09784.
447. Barski A, Cuddapah S, Cui K, et al. High-Resolution Profiling of Histone Methylations in the Human Genome. *Cell*. 2007;129(4):823-837. doi:10.1016/j.cell.2007.05.009.
448. Cao R, Wang L, Wang H, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science*. 2002;298(5595):1039-1043. doi:10.1126/science.1076997.
449. Viré E, Brenner C, Deplus R, et al. The Polycomb group protein EZH2 directly controls DNA methylation. *Nature*. 2005;439(7078):871-874. doi:10.1038/nature04431.
450. Reddington JP, Perricone SM, Nestor CE, et al. Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol*. 2013;14(3):R25. doi:10.1186/gb-2013-14-3-r25.
451. Lindroth AM, Park YJ, McLean CM, et al. Antagonism between DNA and H3K27 Methylation at the Imprinted Rasgrf1 Locus. van Steensel B, ed. *PLoS Genet*. 2008;4(8):e1000145. doi:10.1371/journal.pgen.1000145.
452. Mendenhall EM, Koche RP, Truong T, et al. GC-Rich Sequence Elements Recruit PRC2 in Mammalian ES Cells. Madhani HD, ed. *PLoS Genet*. 2010;6(12):e1001244. doi:10.1371/journal.pgen.1001244.
453. Lynch MD, Smith AJH, De Gobbi M, et al. An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. *EMBO J*. 2012;31(2):317-329. doi:10.1038/emboj.2011.399.
454. Wu H, Coskun V, Tao J, et al. Dnmt3a-Dependent Nonpromoter DNA Methylation Facilitates Transcription of Neurogenic Genes. *Science*. 2010;329(5990):444-448. doi:10.1126/science.1190485.
455. Hagarman JA, Motley MP, Kristjansdottir K, Soloway PD. Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS One*. 2013;8(1):e53880. doi:10.1371/journal.pone.0053880.
456. Neri F, Incarnato D, Krepelova A, et al. Genome-wide analysis identifies a functional association of Tet1 and Polycomb repressive complex 2 in mouse embryonic stem cells. *Genome Biol*. 2013;14(8):R91. doi:10.1186/gb-2013-14-8-r91.
457. Neri F, Incarnato D, Krepelova A, et al. TET1 is controlled by pluripotency-associated factors in ESCs and downmodulated by PRC2 in differentiated cells and tissues. *Nucleic Acids Res*. 2015;43(14):6814-6826. doi:10.1093/nar/gkv392.

458. Harikumar A, Meshorer E. Chromatin remodeling and bivalent histone modifications in embryonic stem cells. *EMBO Rep*. 2015;16(12):1609-1619. doi:10.15252/embr.201541011.
459. Bernstein BE, Mikkelsen TS, Xie X, et al. A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. doi:10.1016/j.cell.2006.02.041.
460. Azuara V, Perry P, Sauer S, et al. Chromatin signatures of pluripotent cell lines. *Nat Cell Biol*. 2006;8(5):532-538. doi:10.1038/ncb1403.
461. Mikkelsen TS, Ku M, Jaffe DB, et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*. 2007;448(7153):553-560. doi:10.1038/nature06008.
462. Marks H, Kalkan T, Menafra R, et al. The transcriptional and epigenomic foundations of ground state pluripotency. *Cell*. 2012;149(3):590-604. doi:10.1016/j.cell.2012.03.026.
463. Rugg-Gunn PJ, Cox BJ, Ralston A, Rossant J. Distinct histone modifications in stem cell lines and tissue lineages from the early mouse embryo. *Proc Natl Acad Sci*. 2010;107(24):10783-10790. doi:10.1073/pnas.0914507107.
464. Sachs M, Onodera C, Blaschke K, Ebata KT, Song JS, Ramalho-Santos M. Bivalent chromatin marks developmental regulatory genes in the mouse embryonic germline in vivo. *Cell Rep*. 2013;3(6):1777-1784. doi:10.1016/j.celrep.2013.04.032.
465. Kong L, Tan L, Lv R, et al. A primary role of TET proteins in establishment and maintenance of *De Novo* bivalency at CpG islands. *Nucleic Acids Res*. 2016;44(18):8682-8692. doi:10.1093/nar/gkw529.
466. Meissner A, Mikkelsen TS, Gu H, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*. 2008;454(7205):nature07107. doi:10.1038/nature07107.
467. Rivas CI, Zúñiga FA, Salas-Burgos A, Mardones L, Ormazabal V, Vera JC. Vitamin C transporters. *J Physiol Biochem*. 2008;64(4):357-375.
468. Daruwala R, Song J, Koh WS, Rumsey SC, Levine M. Cloning and functional characterization of the human sodium-dependent vitamin C transporters hSVCT1 and hSVCT2. *FEBS Lett*. 1999;460(3):480-484.
469. Tsukaguchi H, Tokui T, Mackenzie B, et al. A family of mammalian Na+-dependent L-ascorbic acid transporters. *Nature*. 1999;399(6731):70-75. doi:10.1038/19986.
470. Liang WJ, Johnson D, Jarvis SM. Vitamin C transport systems of mammalian cells. *Mol Membr Biol*. 18(1):87-95.
471. Dhariwal KR, Hartzell WO, Levine M. Ascorbic acid and dehydroascorbic acid measurements in human plasma and serum. *Am J Clin Nutr*. 1991;54(4):712-716.
472. Kim H, Bae S, Yu Y, et al. The analysis of vitamin C concentration in organs of gulo(-/-) mice upon vitamin C withdrawal. *Immune Netw*. 2012;12(1):18-26. doi:10.4110/in.2012.12.1.18.
473. Burns J. Missing Step in Man, Monkey and Guinea Pig required for the Biosynthesis of L-Ascorbic Acid. *Nature*. 1957;180(4585):553-553. doi:10.1038/180553a0.
474. Nishikimisp M, Fukuyaman R, Minoshiman S, Shimizux N, Yagis K. Cloning and Chromosomal Mapping of the Human Nonfunctional Gene for L-Gulono-y-lactone Oxidase, the Enzyme for L-Ascorbic Acid Biosynthesis Missing in Man. *J Biol Chem*. 1994;269:13685-13688.
475. Welch RW, Bergsten P, Butlert JD, Levine M. Ascorbic acid accumulation and transport in human fibroblasts. *Biochem J*. 1993;294:505-510.
476. Braun L, Puskfis F, Csala M, et al. Gluconeogenesis from ascorbic acid" ascorbate recycling in isolated murine hepatocytes. *FIBS 17263 FEBS Lett*. 1996;390:183-186.
477. Zheng Y, Scow JS, Duenes JA, Sarr MG. Mechanisms of Glucose Uptake in Intestinal Cell Lines: Role of GLUT2. doi:10.1016/j.surg.2011.07.010.

478. Drew R, Miners JO. The effects of buthionine sulphoximine (BSO) on glutathione depletion and xenobiotic biotransformation. *Biochem Pharmacol*. 1984;33(19):2989-2994.
479. Meister A. Glutathione-Ascorbic Acid Antioxidant System in Animals. *J Biol Chem*. 1994;269(13):9397-9400.
480. Garcia-Gonzalo FR, Izpisúa Belmonte JC. Albumin-Associated Lipids Regulate Human Embryonic Stem Cell Self-Renewal. Rutherford S, ed. *PLoS One*. 2008;3(1):e1384. doi:10.1371/journal.pone.0001384.
481. Burr S, Caldwell A, Chong M, et al. Oxygen gradients can determine epigenetic asymmetry and cellular differentiation via differential regulation of Tet activity in embryonic stem cells. *Nucleic Acids Res*. November 2017. doi:10.1093/nar/gkx1197.
482. Haffner MC, Pellakuru LG, Ghosh S, et al. Tight correlation of 5-hydroxymethylcytosine and Polycomb marks in health and disease. *Cell Cycle*. 2013;12(12):1835. doi:10.4161/CC.25010.
483. Webb WM, Sanchez RG, Perez G, et al. Dynamic association of epigenetic H3K4me3 and DNA 5hmC marks in the dorsal hippocampus and anterior cingulate cortex following reactivation of a fear memory. *Neurobiol Learn Mem*. 2017;142:66-78. doi:10.1016/j.nlm.2017.02.010.
484. Jones A, Wang H. Polycomb repressive complex 2 in embryonic stem cells: an overview. *Protein Cell*. 2010;1(12):1056-1062. doi:10.1007/s13238-010-0142-7.
485. Boyer LA, Plath K, Zeitlinger J, et al. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature*. 2006;441(7091):349-353. doi:10.1038/nature04733.
486. Cao R, Wang L, Wang H, et al. Role of Histone H3 Lysine 27 Methylation in Polycomb-Group Silencing. *Science*. 2002;298(5595):1039-1043. doi:10.1126/science.1076997.
487. Shen X, Liu Y, Hsu Y-J, et al. EZH1 Mediates Methylation on Histone H3 Lysine 27 and Complements EZH2 in Maintaining Stem Cell Identity and Executing Pluripotency. *Mol Cell*. 2008;32(4):491-502. doi:10.1016/j.molcel.2008.10.016.
488. Kafer GR, Li X, Horii T, et al. 5-Hydroxymethylcytosine Marks Sites of DNA Damage and Promotes Genome Stability. *Cell Rep*. 2016;14(6):1283-1292. doi:10.1016/j.celrep.2016.01.035.
489. Huang K, Fan G. DNA methylation in cell differentiation and reprogramming: an emerging systematic view. *Regen Med*. 2010;5(4):531-544. doi:10.2217/rme.10.35.
490. Hattori N, Nishino K, Ko Y, et al. Epigenetic Control of Mouse *Oct-4* Gene Expression in Embryonic Stem Cells and Trophoblast Stem Cells. *J Biol Chem*. 2004;279(17):17063-17069. doi:10.1074/jbc.M309002200.
491. Farthing CR, Ficz G, Ng RK, et al. Global Mapping of DNA Methylation in Mouse Promoters Reveals Epigenetic Reprogramming of Pluripotency Genes. Frankel WN, ed. *PLoS Genet*. 2008;4(6):e1000116. doi:10.1371/journal.pgen.1000116.
492. Petell CJ, Alabdi L, He M, San Miguel P, Rose R, Gowher H. An epigenetic switch regulates *de novo* DNA methylation at a subset of pluripotency gene enhancers during embryonic stem cell differentiation. *Nucleic Acids Res*. 2016;44(16):7605-7617. doi:10.1093/nar/gkw426.
493. Isagawa T, Nagae G, Shiraki N, et al. DNA Methylation Profiling of Embryonic Stem Cell Differentiation into the Three Germ Layers. Imhof A, ed. *PLoS One*. 2011;6(10):e26052. doi:10.1371/journal.pone.0026052.
494. Fouse SD, Shen Y, Pellegrini M, et al. Promoter CpG methylation contributes to ES cell gene regulation in parallel with Oct4/Nanog, PcG complex, and histone H3 K4/K27 trimethylation. *Cell Stem Cell*. 2008;2(2):160-169. doi:10.1016/j.stem.2007.12.011.
495. Mariani CJ, Vasanthakumar A, Madzo J, et al. TET1-Mediated Hydroxymethylation Facilitates Hypoxic Gene Induction in Neuroblastoma HHS Public Access. *Cell Rep*. 2014;7(5):1343-1352. doi:10.1016/j.celrep.2014.04.040.

496. Wu M-Z, Chen S-F, Nieh S, et al. Hypoxia Drives Breast Tumor Malignancy through a TET-TNFα-p38-MAPK Signaling Axis. *Cancer Res*. 2015;75(18):3912-3924. doi:10.1158/0008-5472.CAN-14-3208.
497. Khoueiry R, Sohni A, Thienpont B, et al. Lineage-specific functions of TET1 in the postimplantation mouse embryo. *Nat Genet*. 2017;49(7):1061-1072. doi:10.1038/ng.3868.
498. Yanes O, Clark J, Wong DM, et al. Metabolic oxidation regulates embryonic stem cell differentiation. *Nat Chem Biol*. 2010;6(6):411-417. doi:10.1038/nchembio.364.
499. Deaton AM, Bird A. CpG islands and the regulation of transcription. *Genes Dev*. 2011;25(10):1010-1022. doi:10.1101/gad.2037511.
500. Zhang H, Zhang X, Clark E, Mulcahey M, Huang S, Shi YG. TET1 is a DNA-binding protein that modulates DNA methylation and gene transcription via hydroxylation of 5-methylcytosine. *Cell Res*. 2010;20(12):1390-1393. doi:10.1038/cr.2010.156.
501. Leitch HG, McEwen KR, Turp A, et al. Naive pluripotency is associated with global DNA hypomethylation. *Nat Struct Mol Biol*. 2013;20(3):311-316. doi:10.1038/nsmb.2510.
502. Ficz G, Hore TA, Santos F, et al. FGF Signaling Inhibition in ESCs Drives Rapid Genome-wide Demethylation to the Epigenetic Ground State of Pluripotency. *Cell Stem Cell*. 2013;13(3):351-359. doi:10.1016/j.stem.2013.06.004.
503. Habibi E, Brinkman AB, Arand J, et al. Whole-Genome Bisulfite Sequencing of Two Distinct Interconvertible DNA Methylomes of Mouse Embryonic Stem Cells. *Cell Stem Cell*. 2013;13(3):360-369. doi:10.1016/j.stem.2013.06.002.
504. Mantsoki A, Devailly G, Joshi A. CpG island erosion, polycomb occupancy and sequence motif enrichment at bivalent promoters in mammalian embryonic stem cells. *Sci Rep*. 2015;5:16791. doi:10.1038/srep16791.
505. Juan AH, Wang S, Ko KD, et al. Roles of H3K27me2 and H3K27me3 Examined during Fate Specification of Embryonic Stem Cells. *Cell Rep*. 2016;17(5):1369-1382. doi:10.1016/j.celrep.2016.09.087.
506. Wu H, Zhang Y. Tet1 and 5-hydroxymethylation: a genome-wide view in mouse embryonic stem cells. *Cell Cycle*. 2011;10(15):2428-2436. doi:10.4161/cc.10.15.16930.
507. Vastenhouw NL, Schier AF. Bivalent histone modifications in early embryogenesis. *Curr Opin Cell Biol*. 2012;24(3):374-386. doi:10.1016/j.ceb.2012.03.009.
508. Tausendschön M, Dehne N, Brüne B. Hypoxia causes epigenetic gene regulation in macrophages by attenuating Jumonji histone demethylase activity. *Cytokine*. 2011;53(2):256-262. doi:10.1016/j.cyto.2010.11.002.
509. Zhou X, Sun H, Chen H, et al. Hypoxia Induces Trimethylated H3 Lysine 4 by Inhibition of JARID1A Demethylase. *Cancer Res*. 2010;70(10):4214-4221. doi:10.1158/0008-5472.CAN-09-2942.
510. Johnson AB, Denko N, Barton MC. Hypoxia induces a novel signature of chromatin modifications and global repression of transcription. *Mutat Res*. 2008;640(1-2):174-179. doi:10.1016/j.mrfmmm.2008.01.001.
511. Shmakova A, Batie M, Druker J, Rocha S. Chromatin and oxygen sensing in the context of JmjC histone demethylases. *Biochem J*. 2014;462:385-395. doi:10.1042/BJ20140754.
512. Niu X, Zhang T, Liao L, et al. The von Hippel–Lindau tumor suppressor protein regulates gene expression and tumor growth through histone demethylase JARID1C. *Oncogene*. 2012;31(6):776-786. doi:10.1038/onc.2011.266.
513. Zhang J, Chen S, Zhang D, et al. Tet3-Mediated DNA Demethylation Contributes to the Direct Conversion of Fibroblast to Functional Neuron. *Cell Rep*. 2016;17(9):2326-2339. doi:10.1016/j.celrep.2016.10.081.
514. Khare T, Pai S, Koncevicius K, et al. 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nat Struct Mol Biol*. 2012;19(10):1037-1043. doi:10.1038/nsmb.2372.

515. Liu N, Wang M, Deng W, et al. Intrinsic and Extrinsic Connections of Tet3 Dioxygenase with CXXC Zinc Finger Modules. Defossez P-A, ed. *PLoS One*. 2013;8(5):e62755. doi:10.1371/journal.pone.0062755.
516. Perera A, Eisen D, Wagner M, et al. TET3 Is Recruited by REST for Context-Specific Hydroxymethylation and Induction of Gene Expression. *Cell Rep*. 2015;11(2):283-294. doi:10.1016/j.celrep.2015.03.020.
517. Weledji EP, Assob JC. The ubiquitous neural cell adhesion molecule (N-CAM). *Ann Med Surg*. 2014;3(3):77-81. doi:10.1016/j.amsu.2014.06.014.
518. Colquitt BM, Allen WE, Barnea G, Lomvardas S. Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity. *Proc Natl Acad Sci U S A*. 2013;110(36):14682-14687. doi:10.1073/pnas.1302759110.
519. Lin Y, Li X-Y, Willis AL, Liu C, Chen G, Weiss SJ. Snail1-dependent control of embryonic stem cell pluripotency and lineage commitment. *Nat Commun*. 2014;5:3070. doi:10.1038/ncomms4070.
520. Nieto MA. The Snail Superfamily of Zinc-Finger Transcription Factors. *Nat Rev Mol Cell Biol*. 2002;3(3):155-166. doi:10.1038/nrm757.
521. Zander MA, Cancino GI, Gridley T, Kaplan DR, Miller FD. The Snail Transcription Factor Regulates the Numbers of Neural Precursor Cells and Newborn Neurons throughout Mammalian Life. Androutsellis-Theotokis A, ed. *PLoS One*. 2014;9(8):e104767. doi:10.1371/journal.pone.0104767.
522. Nordin K, LaBonne C. Sox5 Is a DNA-binding cofactor for BMP R-Smads that directs target specificity during patterning of the early ectoderm. *Dev Cell*. 2014;31(3):374-382. doi:10.1016/j.devcel.2014.10.003.
523. Münzel M, Globisch D, Brückl T, et al. Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain. *Angew Chemie Int Ed*. 2010;49(31):5375-5377. doi:10.1002/anie.201002033.
524. Pfeifer GP, Kadam S, Jin S-G. 5-hydroxymethylcytosine and its potential roles in development and cancer. *Epigenetics Chromatin*. 2013;6(1):10. doi:10.1186/1756-8935-6-10.
525. Jin S-G, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. *Nucleic Acids Res*. 2011;39(12):5015-5024. doi:10.1093/nar/gkr120.
526. Wanunu M, Cohen-Karni D, Johnson RR, et al. Discrimination of Methylcytosine from Hydroxymethylcytosine in DNA Molecules. *J Am Chem Soc*. 2011;133(3):486-492. doi:10.1021/ja107836t.
527. Rodríguez López CM, Lloyd AJ, Leonard K, Wilkinson MJ. Differential Effect of Three Base Modifications on DNA Thermostability Revealed by High Resolution Melting. *Anal Chem*. 2012;84(17):7336-7342. doi:10.1021/ac301459x.
528. Jones-Villeneuve EM, Rudnicki MA, Harris JF, McBurney MW. Retinoic acid-induced neural differentiation of embryonal carcinoma cells. *Mol Cell Biol*. 1983;3(12):2271-2279.
529. McBurney MW, Jones-Villeneuve EM, Edwards MK, Anderson PJ. Control of muscle and neuronal differentiation in a cultured embryonal carcinoma cell line. *Nature*. 1982;299(5879):165-167.
530. Zhang J, Gao Y, Yu M, et al. Retinoic Acid Induces Embryonic Stem Cell Differentiation by Altering Both Encoding RNA and microRNA Expression. Gaetano C, ed. *PLoS One*. 2015;10(7):e0132566. doi:10.1371/journal.pone.0132566.
531. Sharma A, Klein SL, Barboza L, Lodhi N, Toth M. Principles Governing DNA Methylation during Neuronal Lineage and Subtype Specification. *J Neurosci*. 2016;36(5):1711-1722. doi:10.1523/JNEUROSCI.4037-15.2016.
532. Qiao Y, Wang X, Wang R, et al. AF9 promotes hESC neural differentiation through recruiting TET2 to neurodevelopmental gene loci for methylcytosine hydroxylation. *Cell Discov*. 2015;1:15017. doi:10.1038/celldisc.2015.17.
533. Finlay BL, Darlington RB. Linked regularities in the development and evolution of mammalian brains. *Science*. 1995;268(5217):1578-1584.

534. Lopes FM, Schröder R, Júnior MLC da F, et al. Comparison between proliferative and neuron-like SH-SY5Y cells as an in vitro model for Parkinson disease studies. *Brain Res*. 2010;1337:85-94. doi:10.1016/j.brainres.2010.03.102.
535. Påhlman S, Ruusala AI, Abrahamsson L, Mattsson ME, Esscher T. Retinoic acid-induced differentiation of cultured human neuroblastoma cells: a comparison with phorbolester-induced differentiation. *Cell Differ*. 1984;14(2):135-144.
536. Stolt CC, Schlierf A, Lommes P, et al. SoxD Proteins Influence Multiple Stages of Oligodendrocyte Development and Modulate SoxE Protein Function. *Dev Cell*. 2006;11(5):697-709. doi:10.1016/j.devcel.2006.08.011.
537. Lai T, Jabaudon D, Molyneaux BJ, et al. SOX5 Controls the Sequential Generation of Distinct Corticofugal Neuron Subtypes. *Neuron*. 2008;57(2):232-247. doi:10.1016/j.neuron.2007.12.023.
538. Wang Y, Shi J, Chai K, Ying X, Zhou BP. The Role of Snail in EMT and Tumorigenesis. *Curr Cancer Drug Targets*. 2013;13(9):963-972.
539. Kühnöl CD, Würfel C, Staege MS, Kramm C. Snail homolog 1 is involved in epithelial-mesenchymal transition-like processes in human glioblastoma cells. *Oncol Lett*. 2017;13(5):3882-3888. doi:10.3892/ol.2017.5875.
540. Peiró S, Escrivà M, Puig I, et al. Snail1 transcriptional repressor binds to its own promoter and controls its expression. *Nucleic Acids Res*. 2006;34(7):2077-2084. doi:10.1093/nar/gkl141.
541. Leptin M. twist and snail as positive and negative regulators during Drosophila mesoderm development. *Genes Dev*. 1991;5(9):1568-1576.
542. Rembold M, Ciglar L, Yáñez-Cuna JO, et al. A conserved role for Snail as a potentiator of active transcription. *Genes Dev*. 2014;28(2):167-181. doi:10.1101/gad.230953.113.
543. Jorda M, Olmeda D, Vinyals A, et al. Upregulation of MMP-9 in MDCK epithelial cell line in response to expression of the Snail transcription factor. *J Cell Sci*. 2005;118(15):3371-3385. doi:10.1242/jcs.02465.
544. Sun L, Diamond ME, Ottaviano AJ, Joseph MJ, Ananthanarayan V, Munshi HG. Transforming Growth Factor- 1 Promotes Matrix Metalloproteinase-9-Mediated Oral Cancer Invasion through Snail Expression. *Mol Cancer Res*. 2008;6(1):10-20. doi:10.1158/1541-7786.MCR-07-0208.
545. Nishioka K, Kishida T, Masui S, Mazda O. De novo CpG methylation on an artificial chromosome-like vector maintained for a long-term in mammalian cells. *Biotechnol Lett*. 2016;38(4):731-740. doi:10.1007/s10529-015-2029-4.
546. Skiles WM, Kester A, Pryor JH, Westhusin ME, Golding MC, Long CR. Oxygen-induced alterations in the expression of chromatin modifying enzymes and the transcriptional regulation of imprinted genes. *Gene Expr Patterns*. January 2018. doi:10.1016/j.gep.2018.01.001.
547. Lin G, Sun W, Yang Z, Guo J, Liu H, Liang J. Hypoxia induces the expression of TET enzymes in HepG2 cells. *Oncol Lett*. 2017;14(6):6457-6462. doi:10.3892/ol.2017.7063.
548. Beyer S, Kristensen MM, Jensen KS, Johansen JV, Staller P. The Histone Demethylases JMJD1A and JMJD2B Are Transcriptional Targets of Hypoxia-inducible Factor HIF. *J Biol Chem*. 2008;283(52):36542-36552. doi:10.1074/jbc.M804578200.
549. Guo X, Tian Z, Wang X, et al. Regulation of histone demethylase KDM6B by hypoxia-inducible factor-2. *Acta Biochim Biophys Sin (Shanghai)*. 2015;47(2):106-113. doi:10.1093/abbs/gmu122.
550. Semenza GL. Hypoxia-inducible factor 1: master regulator of O2 homeostasis. *Curr Opin Genet Dev*. 1998;8(5):588-594.
551. Wang J, Zhang D, Du J, et al. Tet1 facilitates hypoxia tolerance by stabilizing the HIF-α proteins independent of its methylcytosine dioxygenase activity. *Nucleic Acids Res*. September 2017. doi:10.1093/nar/gkx869.
552. Johansson C, Tumber A, Che K, et al. The roles of Jumonji-type oxygenases in

human disease. *Epigenomics*. 2014;6(1):89-120. doi:10.2217/epi.13.79.
553. Zhang R-R, Cui Q-Y, Murai K, et al. Tet1 regulates adult hippocampal neurogenesis and cognition. *Cell Stem Cell*. 2013;13(2):237-245. doi:10.1016/j.stem.2013.05.006.
554. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315-322. doi:10.1038/nature08514.
555. Zhang J, Chen S, Zhang D, et al. Tet3-Mediated DNA Demethylation Contributes to the Direct Conversion of Fibroblast to Functional Neuron. *Cell Rep*. 2016;17(9):2326-2339. doi:10.1016/j.celrep.2016.10.081.
556. Kang J, Lienhard M, Pastor WA, et al. Simultaneous deletion of the methylcytosine oxidases Tet1 and Tet3 increases transcriptome variability in early embryogenesis. *Proc Natl Acad Sci*. 2015;112(31):E4236-E4245. doi:10.1073/pnas.1510510112.
557. Guo JU, Su Y, Zhong C, Ming G, Song H. Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain. *Cell*. 2011;145(3):423-434. doi:10.1016/j.cell.2011.03.022.
558. Rudenko A, Dawlaty MM, Seo J, et al. Tet1 is critical for neuronal activity-regulated gene expression and memory extinction. *Neuron*. 2013;79(6):1109-1122. doi:10.1016/j.neuron.2013.08.003.
559. Hayashi K, Ohta H, Kurimoto K, Aramaki S, Saitou M. Reconstitution of the Mouse Germ Cell Specification Pathway in Culture by Pluripotent Stem Cells. *Cell*. 2011;146(4):519-532. doi:10.1016/j.cell.2011.06.052.
560. Keller G. Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev*. 2005;19(10):1129-1155. doi:10.1101/gad.1303605.
561. Siggia ED, Warmflash A. Modeling mammalian gastrulation with embryonic stem cells.
562. Harrison SE, Sozen B, Christodoulou N, Kyprianou C, Zernicka-Goetz M. Assembly of embryonic and extraembryonic stem cells to mimic embryogenesis in vitro. *Science*. 2017;356(6334):eaal1810. doi:10.1126/science.aal1810.
563. Warmflash A, Sorre B, Etoc F, Siggia ED, Brivanlou AH. A method to recapitulate early embryonic spatial patterning in human embryonic stem cells. *Nat Methods*. 2014;11(8):847-854. doi:10.1038/nmeth.3016.
564. Hill PWS, Amouroux R, Hajkova P. DNA demethylation, Tet proteins and 5-hydroxymethylcytosine in epigenetic reprogramming: An emerging complex story. *Genomics*. 2014;104(5):324-333. doi:10.1016/J.YGENO.2014.08.012.
565. Gao Y, Chen J, Li K, et al. Replacement of Oct4 by Tet1 during iPSC Induction Reveals an Important Role of DNA Methylation and Hydroxymethylation in Reprogramming. *Cell Stem Cell*. 2013;12(4):453-469. doi:10.1016/j.stem.2013.02.005.
566. Lian X, Zhang J, Azarin SM, et al. Directed cardiomyocyte differentiation from human pluripotent stem cells by modulating Wnt/β-catenin signaling under fully defined conditions. *Nat Protoc*. 2012;8(1):162-175. doi:10.1038/nprot.2012.150.
567. Burridge PW, Zambidis ET. Highly Efficient Directed Differentiation of Human Induced Pluripotent Stem Cells into Cardiomyocytes. In: *Methods in Molecular Biology (Clifton, N.J.)*. Vol 997. ; 2013:149-161. doi:10.1007/978-1-62703-348-0_12.
568. Mummery CL, Zhang J, Ng ES, Elliott DA, Elefanty AG, Kamp TJ. Differentiation of Human Embryonic Stem Cells and Induced Pluripotent Stem Cells to Cardiomyocytes: A Methods Overview. *Circ Res*. 2012;111(3):344-358. doi:10.1161/CIRCRESAHA.110.227512.
569. Bibel M, Richter J, Schrenk K, et al. Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nat Neurosci*. 2004;7(9):1003-1009. doi:10.1038/nn1301.
570. Kim M, Habiba A, Doherty JM, Mills JC, Mercer RW, Huettner JE. Regulation of mouse embryonic stem cell neural differentiation by retinoic acid. *Dev Biol*. 2009;328(2):456-471. doi:10.1016/j.ydbio.2009.02.001.

571. Chuang J-H, Tung L-C, Lin Y. Neural differentiation from embryonic stem cells in vitro: An overview of the signaling pathways. *World J Stem Cells*. 2015;7(2):437-447. doi:10.4252/wjsc.v7.i2.437.
572. Xu X, Tao Y, Gao X, et al. A CRISPR-based approach for targeted DNA demethylation. *Cell Discov*. 2016;2(1):16009. doi:10.1038/celldisc.2016.9.
573. Ellis P, Fagan BM, Magness ST, et al. SOX2, a persistent marker for multipotential neural stem cells derived from embryonic stem cells, the embryo or the adult. *Dev Neurosci*. 2004;26(2-4):148-165. doi:10.1159/000082134.
574. Gusel'nikova V V, Korzhevskiy DE. NeuN As a Neuronal Nuclear Antigen and Neuron Differentiation Marker. *Acta Naturae*. 2015;7(2):42-47.
575. Pataskar A, Jung J, Smialowski P, et al. NeuroD1 reprograms chromatin and transcription factor landscapes to induce the neuronal program. *EMBO J*. 2016;35(1):24-45. doi:10.15252/embj.201591206.
576. Liang H, Hippenmeyer S, Ghashghaei HT. A Nestin-cre transgenic mouse is insufficient for recombination in early embryonic neural progenitors. *Biol Open*. 2012;1(12):1200-1203. doi:10.1242/bio.20122287.
577. Chen J, Kwon C-H, Lin L, Li Y, Parada LF. Inducible site-specific recombination in neural stem/progenitor cells. *Genesis*. 2009;47(2):122-131. doi:10.1002/dvg.20465.
578. Sparrow DB, Chapman G, Smith AJ, et al. A Mechanism for Gene-Environment Interaction in the Etiology of Congenital Scoliosis. *Cell*. 2012;149(2):295-306. doi:10.1016/J.CELL.2012.02.054.
579. Takahashi M, Osumi N. The method of rodent whole embryo culture using the rotator-type bottle culture system. *J Vis Exp*. 2010;(42). doi:10.3791/2170.
580. Feil D, Lane M, Roberts CT, et al. Effect of culturing mouse embryos under different oxygen concentrations on subsequent fetal and placental development. *J Physiol*. 2006;572(Pt 1):87-96. doi:10.1113/jphysiol.2005.102681.
581. Bachman M, Uribe-Lewis S, Yang X, et al. 5-Formylcytosine can be a stable DNA modification in mammals. *Nat Chem Biol*. 2015;11(8):555-557. doi:10.1038/nchembio.1848.
582. Raiber E-A, Beraldi D, Ficz G, et al. Genome-wide distribution of 5-formylcytosine in embryonic stem cells is associated with transcription and depends on thymine DNA glycosylase. *Genome Biol*. 2012;13(8):R69. doi:10.1186/gb-2012-13-8-r69.
583. Fu Y, Dominissini D, Rechavi G, He C. Gene expression regulation mediated through reversible m6A RNA methylation. *Nat Rev Genet*. 2014;15(5):293-306. doi:10.1038/nrg3724.
584. Motorin Y, Lyko F, Helm M. 5-methylcytosine in RNA: detection, enzymatic formation and biological functions. *Nucleic Acids Res*. 2010;38(5):1415-1430. doi:10.1093/nar/gkp1117.
585. Fu L, Guerrero CR, Zhong N, et al. Tet-mediated formation of 5-hydroxymethylcytosine in RNA. *J Am Chem Soc*. 2014;136(33):11582-11585. doi:10.1021/ja505305z.
586. Fedeles BI, Singh V, Delaney JC, Li D, Essigmann JM. The AlkB Family of Fe(II)/α-Ketoglutarate-dependent Dioxygenases: Repairing Nucleic Acid Alkylation Damage and Beyond. 2015. doi:10.1074/jbc.R115.656462.
587. Fu Y, Jia G, Pang X, et al. FTO-mediated formation of N6-hydroxymethyladenosine and N6-formyladenosine in mammalian RNA. *Nat Commun*. 2013;4:1798. doi:10.1038/ncomms2822.
588. Li D, Zhang B, Xing X, Wang T. Combining MeDIP-seq and MRE-seq to investigate genome-wide CpG methylation. *Methods*. 2015;72:29-40. doi:10.1016/j.ymeth.2014.10.032.